

Structural and functional characterization of the core splicing factor PRP8

by

Garima Mehta

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Biochemistry

University of Alberta

© Garima Mehta, 2015

***Abstract:***

Over 90% of eukaryotic genes are initially expressed as precursor-messenger RNAs (pre-mRNAs) that contain coding exon sequences interrupted by non-coding intron sequences. The introns are excised and exons are ligated together to form messenger RNA (mRNA) through a process called splicing. Splicing involves two sequential transesterification steps catalyzed by a large RNA/protein complex called the spliceosome.

A number of studies implicate a protein called Prp8 at the heart of the catalytic core of the spliceosome. I am interested in determining the exact role of the RNase H domain of Prp8 in regulating the individual steps of splicing and the transition between the two steps of splicing. In order to understand the function of Prp8 in this regard, I have created panels of mutant RNase H Prp8 proteins that are predicted to affect either the switch between Prp8 first step and second step conformations or the ability of Prp8 to bind a metal ion in the second step conformation.

The human Prp8 mutant alleles will be characterized structurally by X-ray methods and the corresponding yeast Prp8 alleles will be tested for their splicing activity both within yeast and biochemical extracts prepared from yeast. In particular, our model suggests switch dependent interactions of the RNase H domain with the rest of Prp8 and these experiments will test this. The experiments with respect to metal binding are important because they will provide more definitive information regarding the role of Prp8 in the chemistry of the splicing reaction. These results will be of great interest

because it has long been believed that RNA, rather than protein components, of the spliceosome were solely responsible for promoting the splicing chemistry.

## ***Preface***

This thesis is an original work by Garima Mehta. No part of this thesis has been previously published.

## *Acknowledgements*

I would like to thank everyone who has been a part of this journey with me. I would like to thank my Parents who have always encouraged and supported me in all my endeavors and other special people in my life, you know who you are, for helping me survive the stress and their unwavering support at all times.

I would like to specially thank my supervisor, Dr. Andrew MacMillan, for giving me the opportunity to work in his lab, his invaluable guidance and encouragement throughout my degree. To my lab members Erin Garside and Matthew Schellenberg, who have collaborated with me for the work presented here. To my other lab members for being cooperative and understanding. To my committee members, Dr. J.Lemieux, Dr.M.Glover and Dr.P.LaPointe for their expert guidance.

Thank you to everyone who I may have missed mentioning unintentionally.

## ***Table of Contents:***

Abstract	<i>ii</i>
Preface	<i>iv</i>
Acknowledgements	<i>v</i>
Table of Contents	<i>vi</i>
List of Figures	<i>viii</i>
List of Abbreviations	<i>ix</i>
Chapter 1. Introduction. Pre-mRNA splicing and spliceosome assembly	1
1-1 Pre-mRNA splicing	2
1-2 Spliceosome assembly	2
1-3. U5 snRNP	8
1-4. Prp8 - an essential spliceosomal protein	9
1-5. Domain organization of Prp8	10
1-5.1. N-terminal region	11
1-5.2. Central region	11
1-5.3. C-terminal region	11
1-6. Prp8-RNA interactions	13
1-7. Prp8-Protein interactions	14
1-8. Conformational change in Prp8 between the first and second step of splicing	14
1-9. Crystal structure of Prp8 RNase H domain	16
1-10. Summary	20
Chapter 2. Materials and methods	21
2-1.1. Protein expression and purification	22
2-1.2. Crystallization	22
2-1.3. Data collection and processing	22
2-1.4. Model building and refinement	23
2-1.5. Construction of RNA substrates	23
2-1.6. Creation of mutant <i>S. cerevisiae</i> strains and splicing extracts	23

2-1.7. Growth Assays	24
2-4.8. Primer extension	25
Chapter 3. Local Regulation of the metal binding site	27
3-1. Introduction	28
3-2. Results	30
3-2.1. Copper growth assay	30
3-2.2. Primer extension assay	31
3-2.3. Structural analysis	33
3-3. Discussion	36
3-4. Summary	37
Chapter 4. Local regulation of the $\beta$ -hairpin/ loop dynamics	38
4-1. Introduction	39
4-2. Results	42
4-2.1. Mutations of L1798 residue regulate the dynamics of the $\beta$ -hairpin region	42
4-2.2. Mutations of N1797 residue regulates the dynamics of the $\beta$ -hairpin region	45
4-3. Discussion	49
4-4. Summary	50
Chapter 5. Conclusions and future directions	52
5-1. Conclusions	53
5-2. Future directions	55
5-2.1. Designing new mutant alleles for the RNase H domain	55
5-2.2. Observe metal binding in Prp8 RNase H domain of <i>S. cerevisiae</i>	57
5-3. Summary	57
References	58

## ***List of Figures:***

### Chapter 1.

I-1. Representation of the two-step splicing mechanism	4
I-2. Stepwise assembly of the spliceosome	7
I-3. Structure of yeast Prp8 (residues 885-1824)	12
I-4. Spliceosomal transition between the two steps of splicing	16
I-5. X-ray structure of the PRP8 RH domain	18
I-6. Detail of coordination of Mg <sup>2+</sup> ion	19

### Chapter 3

3-1. Superimposition of the X-ray structure of the Prp8 RNase H domain two conformations.	29
3-2. Growth phenotypes of mutant strains with ACT1-CUP1-BSG reporter system.	31
3-3. Denaturing PAGE analysis of reverse transcriptase primer extension.	34
3-4. Characterization of Prp8 RNase H domain residue T1781.	35

### Chapter 4.

4-1. Mapping of yeast mutant alleles on the Prp8 RNase H domain.	39
4-2. Models for the L1798 mutants.	41
4-3. Growth phenotypes of mutant strains with ACT1-CUP1-BSG reporter system.	42
4-4. Denaturing PAGE analysis of reverse transcriptase primer extension of L1798 mutants.	43
4-5. Structure of the Prp8 second step allele L1798N.	45
4-6. Structure of the Prp8 second step allele N1797D.	46
4-7. Denaturing PAGE analysis of reverse transcriptase primer extension N1797D mutant.	48

### Chapter 5.

5 Model representing the role of Prp8 RNase H domain in splicing.	54
---	----

## ***List of Tables:***

1-1. Major Protein components of snRNPs	5
---	---

***List of Abbreviations:***

5' SS	5' splice site
3' SS	3' splice site
5-FOA	5-fluoroorotic acid
A	adenosine
aa	amino acid
ATP	adenosine triphosphate
BPS	branch point sequence
BSC	ACT-1 pre-mRNA branch site A to C mutation
BSG	ACT-1 pre-mRNA branch site A to G mutation
C	cytosine
CC	commitment complex
E complex	early complex
G	guanidine
MBP	maltose-binding protein
mRNA	messenger RNA
nt	nucleotide
PCR	polymerase chain reaction
PPT	poly-pyrimidine tract
PROCN	Prp8 central domain
PRP	pre-mRNA processing
Pre-mRNA	precursor messenger-RNA
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein particle
ss	splice site
TERT	telomerase reverse transcriptase
Th/X	thumb/X/maturase domain
U2AF	U2 auxiliary factor

**Note: Yeast Prp8 amino acid numbers differ from human Prp8 amino acid numbers by 72 (e.g: T1855 in yPrp8 = T1783 in hPrp8)**

## **Introduction**

### **Pre-mRNA splicing and spliceosome assembly**

### **1-1. Pre-mRNA splicing**

Eukaryotic pre-mRNAs are characterized by a split gene structure in which the protein coding exon sequences are interrupted by non-coding intron sequences. The split gene structure was first discovered during the mapping of adenoviral gene structure where it was seen that the hybridization between the isolated messenger RNA (mRNA) of adenovirus-2 and its encoding DNA resulted in the formation of un-hybridized loops at certain intervals (Berget *et al.* 1977; Chow *et al.* 1977). Later it was discovered that these loops correspond to introns. The exons hybridize to their complementary sequences in the mRNA, resulting in loop formation of un-hybridized introns. Such an observation led to the realization that introns are removed from the pre-mRNA before forming the mRNA.

During gene expression in the majority of eukaryotes, DNA is first transcribed to a precursor-messenger RNA (pre-mRNA) by RNA polymerase. The process by which introns are excised from the pre-mRNA and the exons are ligated together to produce a mRNA which encodes a functional protein is referred to as pre-mRNA splicing. Pre-mRNA splicing takes place through two sequential phosphotransesterification reactions that cleave the pre-mRNA, excise the intron and join the exons. Pre-mRNA splicing is a crucial step in gene expression because it is required for protein production in eukaryotes. Mutations that affect splicing can directly cause a disease or contribute to the susceptibility of a disease.

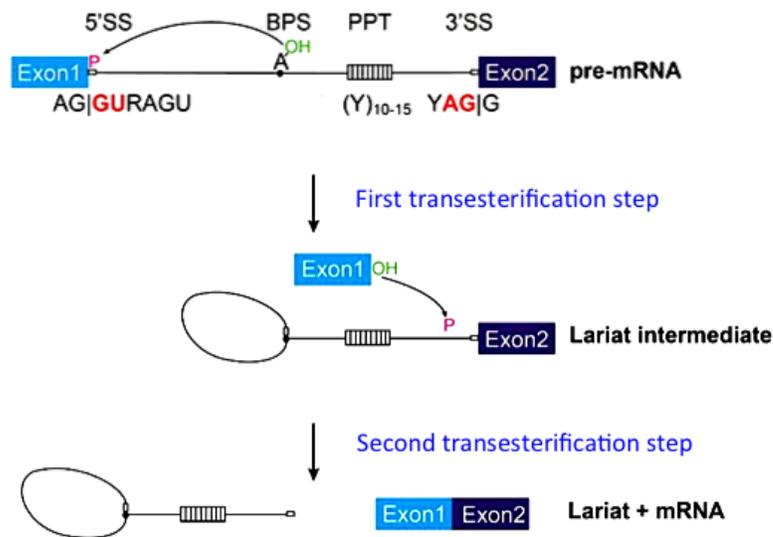
### **1-2. Spliceosome assembly**

Four different types of RNA splicing have been identified. Group I and group II RNA splicing is self-catalyzed. Group I splicing requires the binding of a guanine (G)

nucleotide to the intron sequence, which is used as a nucleophile to attack the 5' splice site (5' SS). The newly formed 3' end of the mRNA then attacks the 3' splice site (3' SS) to complete the reaction (reviewed in: Nielsen and Johansen, 2009). Group II introns have a similar mechanism of splicing, but the nucleophile in the first step is an adenosine (A) residue within the intron. (reviewed in: Ritchie et al., 2009). Group III splicing reactions, found in nuclear mRNA primary transcripts, occur within the spliceosome which is a large cellular complex containing both RNA and protein components (Rio 1993; Valadkhan and Jaladat, 2010). Group II introns and nuclear mRNA primary transcripts share similar transesterification steps, but differ significantly in consensus sequence and snRNP requirement. The group IV class of splicing is different from the above three groups in that the splicing reaction requires ATP and a protein endonuclease (reviewed in: Calvin and Li, 2008).

The spliceosome recognizes the Group III spliceosomal introns defined by short, conserved sequences. The 5' SS contains a universally conserved GU dinucleotide at the 5' end of the intron. The 3' end of the intron is comprised of a seven nucleotide branch point sequence (BPS) followed by a downstream poly-pyrimidine (cytosine and uracil) tract (PPT). Metazoans have a divergent consensus sequence at the branch point 'YNCURAC' (Y=pyrimidine, R=purine and N=any nucleotide, A=branch adenosine ) while the yeast, *Saccharomyces cerevisiae*, has the conserved consensus sequence of 'UACUAAC' as its BPS. A universally conserved AG dinucleotide marks the 3' end (3' SS) of the intron. Many splicing enhancer and repressor sequences in the vicinity of splice sites are present to modulate splicing in metazoans (Barash *et al.*, 2010).

Once the spliceosome has identified an intron within a pre-mRNA, it initiates the two phosphotransesterification reactions to remove the intron. The first transesterification reaction involves an attack by the 2' hydroxyl (2' OH) of a conserved adenosine at the BPS at the 5' SS resulting in the formation of a branched lariat intermediate containing a 2'-5' phosphodiester linkage and a free 5' exon. Next, the 3' OH of the free 5' exon attacks the 3' SS via a second transesterification reaction resulting in a lariat-structure intron and ligated exons (Moore *et al.* 1993) (Figure I-1).



**Figure 1-1.** Representation of the two-step splicing mechanism. Critical sequence elements are also shown. Y =pyrimidine, R=purine.

The spliceosome is composed of more than three hundred protein factors and five small nuclear RNAs (snRNAs), namely U1, U2, U4, U5 and U6 snRNA. snRNAs are believed to have evolved from group II introns since all snRNAs, with the exception of U4 snRNA, have functional counterparts in group II introns. (Cech, 1986; Valadkhan, 2010).

Intron recognition and removal takes place in an ordered, step-wise assembly of small nuclear ribonucleoprotein particles (snRNPs). snRNPs are protein-RNA complexes named for the specific snRNA associated with it : U1, U2, U4, U5, and U6 snRNP. Each snRNP contains seven common core proteins called Smith (Sm) proteins (SmB/SmB', SmD1, SmD2, SmD3, SmE, SmF, and SmG) or Sm-like proteins (in case of U6) named after the patient whose serum contained antibodies specific for the Sm complex (Kattah *et al.*, 2010). snRNPs also contain some snRNP-specific proteins such as Prp8, Brr2, Snu114, to name a few, in the U5 snRNP (Will and Luhrmann 2001a) (Table I-1).

**Table 1-1. Major Protein components of snRNPs**

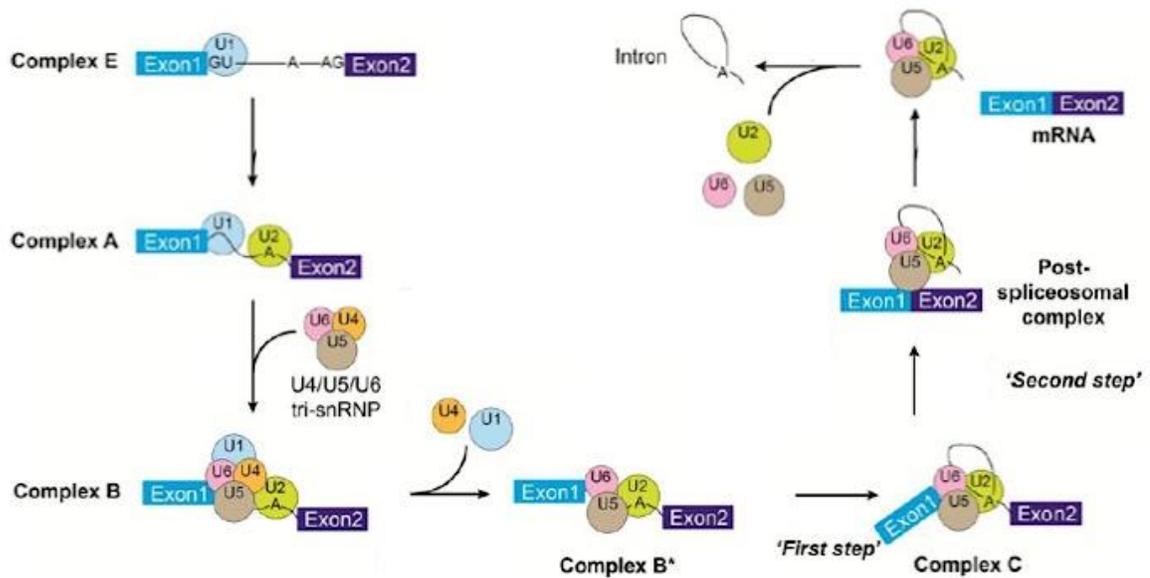
U1	U2	U4	U5	U6
Sm	Sm	Sm	Sm	Lsm
U1A	U2A'	Prp3	Prp8	Prp24
U1C	U2B''	Prp4	Brr2	
U170K	Sf3a	Prp31	Snu114	
Prp40	Sf3b	CypH	Prp6	
	Prp9	Snu13	Prp28	
		15.5K	Snu40	
			Dib1	

The spliceosome is re-assembled anew upon each pre-mRNA. Spliceosome assembly is initiated by the ATP-independent binding of the U1 snRNP to the 5' SS to form the commitment complex (CC) in yeast or early complex (E) in mammals on the pre-mRNA. This binding is mediated by a base pairing interaction between the 5' SS and the U1 snRNA (Query *et al.* 1994). U1 snRNP associated proteins assist in the interaction between the 5' SS and U1 snRNA (Zhang & Rosbash, 1999). Other non-snRNP associated proteins, such as U2 snRNP auxiliary factor (U2AF) and splicing factor 1 (SF1) recognize the PPT/3' SS and BPS respectively (Wu *et al.*, 1999; Zorio *et al.*, 1999; Merendino *et al.*, 1999). The formation of this complex commits the substrate to the splicing pathway. U2AF leads to the recruitment of U2 snRNP in an ATP-dependent manner to form the pre-spliceosome or A complex. Imperfect base pairing between U2 snRNA and the pre-mRNA branch region extrudes the branch adenosine, selecting it as the nucleophile for the first transesterification reaction of splicing (Query *et al.* 1994).

Next, the U4/U6•U5 tri-snRNP is recruited to the 5' SS through interactions with U1 snRNP to form the precatalytic spliceosome or B complex. Extensive base pairing between U4 and U6 snRNA in the U4/U6•U5 tri-snRNP is disrupted to allow U2/U6 snRNA base pairing (Staley and Guthrie 1999). The U2/U6 snRNA base pairing has been shown to be essential for catalysis of the splicing transesterification reactions (Madhani and Guthrie, 1992; Brow, 2002; Valadkhan *et al.*, 2007). Additionally, the U1/5' SS binding is replaced by the base-pairing interaction between the U6 snRNA (ACAGAGA box) and the 5' SS. This change is mediated by the helicase Prp28 (Staley and Guthrie, 1999). The U1 and U4 snRNPs are released from the spliceosome, the

nineteen complex (NTC) is formed, which includes Prp19 protein contributing to stabilizing the U5 and U6 snRNP binding to the spliceosome (Chan *et al.*, 2003).

Rearrangement of RNA/RNA and RNA/protein structures in the spliceosome drive the spliceosome into the catalytic step one spliceosome or C complex. The first transesterification step of splicing occurs in the C complex. Further rearrangements take place in the C complex before the second catalytic step. After the second transesterification step occurs and the exons have been ligated, the spliceosome dissociates and the mature mRNA is released (Figure 1-2). snRNPs are recycled for the next round of splicing.



**Figure 1-2.** Stepwise assembly of the spliceosome on a pre-mRNA showing sequential association of U1, U2, U4/U6•U5 snRNPs with the intron. Shown are the formation of the E (commitment), A, and B complexes through to the mature spliceosome in C complex. (Adapted from Will *et al.*, 2010).

Two additional RNA helicases, namely Prp16 and Prp22, are required for splicing after the spliceosome is fully assembled. Helicase Prp16 is involved in proofreading the first step of splicing by discarding sub-optimal substrates. Prp16 also causes a conformational change of the spliceosome required for the second step of splicing (Lardelli *et al.*, 2010; Schwer and Guthrie 1992). RNA helicase Prp22 is involved in proofreading the second step of splicing (exon ligation) by detecting aberrant substrates. Prp22, in combination with Brr2, also helps to release the mature mRNA (Schwer, 2008). Another helicase, Prp43, is involved in discarding the defective substrates, release of the spliced intron and disassembly of the spliceosome once the splicing reactions are completed (Pandit *et al.*, 2006; Koodathingal *et al.*, 2010).

Self-splicing group II introns and the spliceosome have structural and mechanistic similarities leading to the hypothesis that spliceosomes are also ribozymes. A study was conducted where a protein-free system containing only the U2/U6 snRNA and two short RNA substrates was able to carry out the second step of splicing indicating that the pre-mRNA splicing is intrinsically RNA catalyzed (Valadkhan *et al.*, 2009). However, crosslinking studies have shown two proteins, namely p14 and Prp8, within 15Å of the branch adenosine (MacMillan *et al.*, 1994, Query *et al.*, 1996). Hence, further studies to understand the function of Prp8 and other spliceosomal proteins will enhance our understanding of pre-mRNA splicing.

### **1-3. U5 snRNP**

Prp8 forms a complex with U5 snRNA, seven Sm core proteins, Snu114, and Aar2 in the cytoplasm to form the U5 snRNP. Once the U5 snRNP is imported into the nucleus, Brr2 protein replaces Aar2 to form the mature U5 snRNP. Brr2, a DExD/H-box

family helicase, the EF2-like GTPase Snu114, and Prp8 are important for the activation of the spliceosome (Wahl *et al.*, 2009). Brr2 plays a role in the release of the mature mRNA from the spliceosome and recycling of the U2, U5 and U6 snRNPs for another round of splicing. Snu114 regulates the activity of Brr2 in a nucleotide dependent manner (Small *et al.*, 2006).

#### **1-4. Prp8 - an essential spliceosomal protein**

Some studies suggest that splicing is catalyzed by RNA, making the spliceosome a ribozyme (Nissen *et al.*, 2000). However, interaction between protein factors and the active site of the spliceosome suggest a cooperative role between snRNAs and the spliceosomal proteins. There is evidence to suggest an interaction between the spliceosomal proteins and the catalytic center of the spliceosome (Collins and Guthrie, 2000). One such spliceosomal protein is Prp8, which has been shown to interact with the 5' SS, 3' SS and the BPS. Additionally, crosslinking experiments give evidence for direct interaction between Prp8 and U6 and U5 snRNAs. These observations suggest that the spliceosome active site may include proteins as well as RNA components, supporting the idea that the catalytic heart of the spliceosome may feature proteins. Additionally, a number of mutant Prp8 alleles suggest an interaction between Prp8 and substrate pre-mRNA and also with snRNA in the spliceosome (Grainger and Beggs, 2005).

Prp8 is a component of the U5 snRNP and therefore a part of the U4/U6-U5 tri-snRNP (Lossky *et al.*, 1987; Gottschalk *et al.*, 1999; Stevens *et al.*, 2001). Prp8 was first identified in a screen for temperature sensitive mutants in *Saccharomyces cerevisiae* and was named such to indicate its involvement in pre-mRNA processing (Hartwell, 1967; Hartwell *et al.*, 1970; Vijayraghavan *et al.*, 1989). A >60% sequence identity between *S.*

*cerevisiae* and human PRP8 genes makes it one of the most highly conserved proteins known. Owing to its large size varying from ~230 kD to ~280 kD in different organisms, it is also one of the largest known spliceosomal factors.

Lethal yeast Prp8 knockdown has established the functional importance of the Prp8 gene. A study in *C. elegans* has shown that injecting siRNA corresponding to Prp8 arrests embryogenesis at the late-gastrulation stage (Takahashi *et al.*, 2001). A severe form of Retinitis Pigmentosa, a genetic condition causing blindness in humans, is due to mutations within the C-terminus of Prp8 protein (Tiegelkamp *et al.*, 1995). Studies have shown that Prp8 plays a critical role in the process of splicing, suggesting that Prp8 could potentially act as a splicing cofactor. It may do so by stabilizing other components within the spliceosome or Prp8 could itself be involved directly in one or both steps of splicing (Abelson, 2008).

For the remainder of my thesis, human Prp8 will be referred to as hPrp8 while yeast Prp8 will be called yPrp8. If a statement is equally relevant to both homologues, Prp8 will be used.

### **1-5. Domain organization of Prp8**

The primary amino acid sequence of Prp8 does not exhibit any obvious conserved protein motifs, making it difficult to predict the domains of this protein. The most recent organizational model based on sequence alignment of Prp8 divides it into three domains, namely N-terminal, central and C-terminal regions (Figure 1-3).

### **1-5.1. N-terminal region**

The N-terminal region can be further divided into three domains. The proline-rich tract or P region is found in all fungal Prp8 proteins and is absent in most other organisms. Next, a nuclear localization signal (NLS) region present downstream of the P region has been shown to mediate the transportation of Prp8 from the cytoplasm into the nucleus (Boon et al., 2007). Lastly, a bromodomain or Br domain region that has several residues those are important for protein-protein interactions and structure stabilization (Dlakic et al., 2011).

### **1-5.2. Central region**

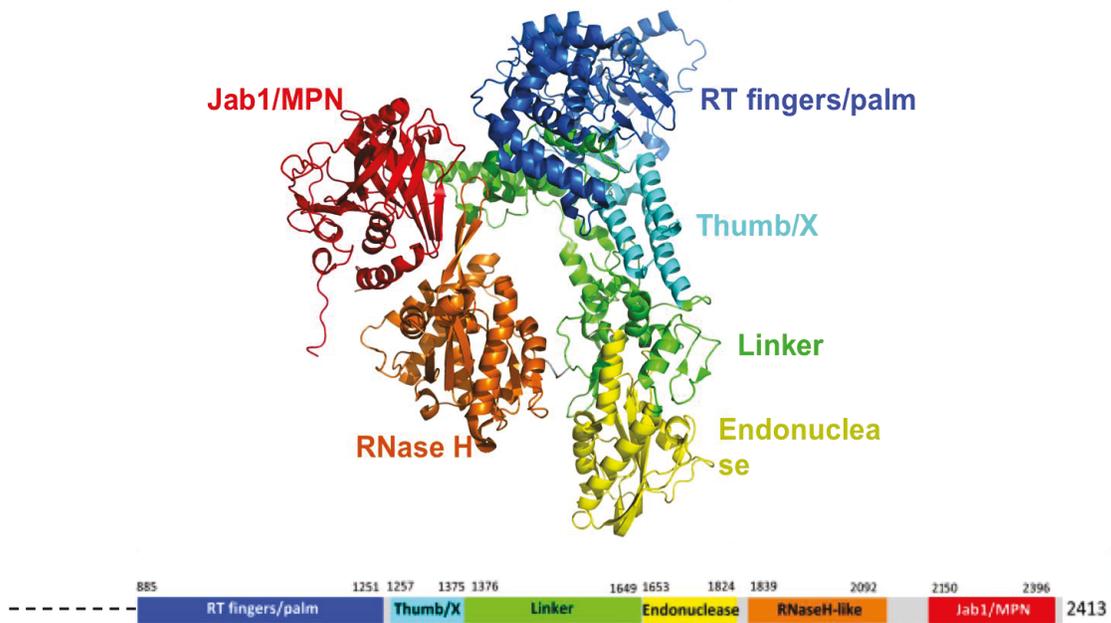
Studies have shown that this region can be sub-divided into a large polymerase-like domain and a small type II restriction endonuclease-like domain that interact with each other through a linker domain (Galej *et al.*, 2013). Sequence similarity between the central part of Prp8 and a catalytic domain of reverse transcriptase from bacterial group II intron encoded proteins have been reported earlier (Dlakic and Mushegian, 2011). The type II restriction endonuclease-like fold has the characteristic five mixed  $\beta$ -strands flanked by three  $\alpha$ -helices.

### **1-5.3. C-terminal region**

The C-terminal region is the best studied region in Prp8. It consists of an RNase H domain (originally named domain IV) and a Jab1/MPN domain. The RNase H domain houses a number of Prp8 alleles that have been shown to suppress mutations on the 5' SS, 3' SS, branch point, poly-pyrimidine tract, U4 snRNA and U6 snRNA (Liu *et al.*, 2007; Grainger and Beggs, 2005; Umen and Guthrie, 1995a; Umen and Guthrie, 1995b; Kuhn

*et al.*, 1999). A stretch of five amino acids in hPrp8 RNase H domain (QACLK, h1894-1898) have shown to cross-link to the 5' SS (Reyes *et al.*, 1999).

The Jab1/MPN domain marks the end of the C-terminal region of Prp8. It is proposed to mediate protein-protein interaction between Prp8 and other U5 snRNP proteins such as the helicase Brr2. The Jab/MPN domain has been shown to stimulate the helicase activity of Brr2 (Maeder *et al.*, 2009; Mozaffari-Jovin *et al.*, 2012). Several mutations within the Jab1/MPN domain of Prp8 have been identified in retinitis Pigmentosa (McKie *et al.*, 2001; Towns *et al.*, 2010).



**Figure 1-3.** Structure of yeast Prp8 (residues 885-1824). RT fingers/palm: reverse transcriptase-like domain; Th/X: conserved domain in Prp8 and a subset of fungal RT-like proteins, located at the same position as “maturase-specific” X/thumb domain; Linker domain; Endonuclease domain: type II restriction endonuclease-like domain; RNase H domain; Jab/MPN domain (Galej *et al.*, 2013).

## 1-6. Prp8-RNA interactions

Presently, Prp8 is the only spliceosomal protein that is known to make extensive contacts with all the reactive sites of the pre-mRNA that are considered to be in or proximal to the catalytic center of the spliceosome. Additionally, Prp8 has also shown to interact with several protein components of other snRNPs. These interactions make Prp8 a leading candidate for a protein cofactor directly contributing to splicing catalysis. Crosslinking studies in yeast and HeLa nuclear extracts show direct interaction between Prp8 and the 5' SS, 3' SS, BPS and PPT (Wyatt *et al.*, 1992; MacMillan *et al.*, 1994; Teigelkamp *et al.*, 1995; Umen and Guthrie a, 1995; Reyes *et al.*, 1996). The interaction between Prp8 and the 5' SS has been shown to occur in the B complex before the first step of splicing, while Prp8 interaction with BPS and 3' SS occurs concurrently or subsequent to the first step. Prp8 has also been shown to interact with U5 (Dix *et al.*, 1998) and U6 snRNAs (Vidal *et al.*, 1999). A number of Prp8 mutants have shown to suppress the mutations of the 5' SS, 3' SS, BPS, PPT, U4 snRNA and U6 snRNA (Grainger and Beggs, 2005).

Proteolytic methods have mapped amino acids QACLK (aa 1894-1898) near the C-terminus of hPrp8 cross-linked to the 5' SS (Reyes *et al.*, 1999). This interaction occurs when Prp8 is a part of the U4/U6•U5 tri-snRNP (Maroney *et al.*, 2000). The Newman lab conducted an experiment where they randomly inserted tobacco etch virus (TEV) protease sites throughout the PRP8 gene to map the crosslinks between Prp8 and the spliceosomal RNA (Turner *et al.*, 2006). In this TEV insertion assay, it was shown that that the Prp8 region 871-970 contacts the 5' SS and a residue 2 nucleotides downstream of the intron branch site. The Newman lab has also mapped the interaction

between aa 1281-1413 of Prp8 with the 5' SS, BPS and the invariant U-rich loop (loop1) of U5 snRNA while amino acids 1503-1673 of Prp8 crosslink to nucleotide 54 in U6 snRNA which is present immediately downstream of the conserved ACAGAG motif that contacts the 5' SS (Vidal *et al.*, 1999; Turner *et al.*, 2006).

At the 3' end of intron, the cross-linking between Prp8 and pre-mRNA spans a region from the BPS to 3' SS and includes bases in the 3' exon. This binding is not specific to the 3' SS sequence (McPheeters and Muhlenkamp, 2003). It has been proposed that the association of Prp8 with the 3' SS results in conformational changes which lock the free 5' exon into the spliceosome (Umen and Guthrie, 1995a).

### **1-7. Prp8-Protein interactions**

Prp8 has been shown to form complexes with other U5 snRNP protein components, namely Brr2 and Snu114 as indicated by yeast two-hybrid (YTH) screens and affinity purification pull-down assays (Dix *et al.*, 1998). U1 snRNP proteins Prp39 and Prp40 have been shown to interact with the N-terminus of Prp8 (van Nues and Beggs, 2001, Abovich and Rosbash, 1997).

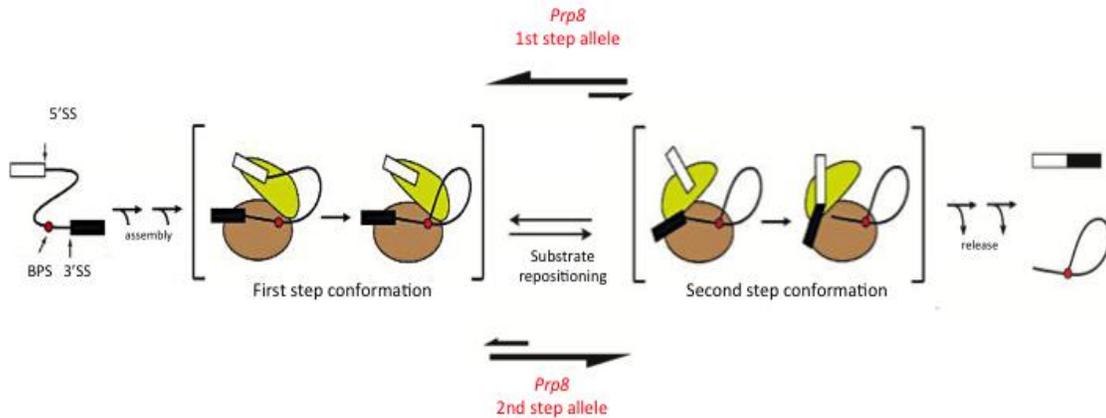
### **1-8. Conformational change in Prp8 between the first and second step of splicing**

Splicing takes place through two consecutive transesterification steps involving three reactive sites within an intron. The first step involves the attack by the branch adenosine on the 5' splice site (5' SS), yielding a lariat intermediate and a cleaved 5' exon. During the second transesterification step, the 5' exon attacks the 3' splice site (3' SS), producing ligated mRNA and lariat intron (reviewed in: Konarska and Query, 2005). Since the substrates for the two transesterification steps are different, rearrangements of

either the substrates or the spliceosome are required at the catalytic center. If the spliceosome uses a single active site for both the catalytic steps of splicing, the lariat intermediate product of the first step would need to be displaced to allow positioning of the 3' SS for the second step (Steitz and Steitz, 1993).

The stereochemical analysis of splicing indicates that there are two distinct active sites or one active site in two conformations which could be formed by the rearrangement of the spliceosome (Moore and Sharp, 1993). In this case, the spliceosome would exist in two distinct conformations during catalysis by binding the substrates differently during the two transesterification steps.

A previous study has proposed a model whereby Prp8 is involved in the equilibrium between two distinct spliceosomal conformations which are associated with the first and second transesterification steps of splicing (Konarska and Query, 2005) (Figure 1-4). This study also proposes that there are two sets of Prp8 alleles, which can shift the equilibrium to favor one step over the other. Thus, a first step allele stabilizes the first step conformation and suppresses a defect in the first but enhances a defect in the second step of splicing, whereas a second step allele stabilizes the second step conformation and suppresses a defect in the second but enhances a defect in the first step of splicing.



**Figure 1-4.** The first and second catalytic steps require different conformations of the spliceosome shown by green and brown ovals. The equilibrium between these conformations is modulated by Prp8, among other protein factors. As indicated by the arrows, certain alleles of Prp8 improve the first step and inhibit the second efficiency, where other alleles act oppositely (Query and Konarska, 2004).

### 1-9. Crystal structure of Prp8 RNase H domain

The presence of a large number of alleles in the RNase H domain that suppress mutations at various sites within the pre-mRNA and U6 snRNA, make the RNase H domain an important player with respect to spliceosome assembly (reviewed in: Grainger and Beggs, 2005). RNase H domains are known to be metalloenzymes promoting hydrolysis of an RNA substrate. Thus, the existence of the RNase H domain in Prp8 and its contact with the splice sites and U6 snRNA make it a good candidate to carry out the role of catalysis within the active site of spliceosome.

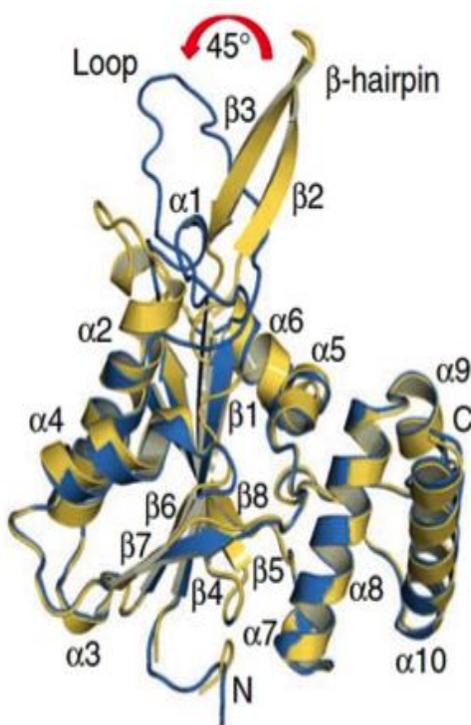
The structure of the RNase H domain of hPrp8 (amino acids 1769-1990) shows that the Prp8 RNase H domain core is bipartite. It consists of an N-terminal sub-domain (amino acids 1769-1887) harboring an RNase H fold and a cluster of five helices making up the C-terminal domain (amino acids 1918-1990). The RNase H fold is characterized

by a five-stranded mixed parallel/anti-parallel  $\beta$ -sheet buttressed by two  $\alpha$ -helices (Ritchie *et al.*, 2008). The RNase H fold was not predicted by primary sequence analysis because it is interrupted by a seventeen amino acid insertion (amino acids 1787-1803) between two adjacent  $\beta$ -strands ( $\beta$ 2/ $\beta$ 3). The hPrp8 structure includes two monomers in the asymmetric unit. In one monomer, monomer a or the closed conformation, this seventeen amino acid insertion is well structured and forms a two-stranded anti-parallel  $\beta$ -hairpin. In the other monomer, monomer b or the open conformation, this hairpin is disordered to form a displaced loop. The loop is translated back by  $\sim 45^\circ$ , pulling aa 1782-1784 to extend the  $\beta$ 1 strand of the RNase H fold (Schellenberg *et al.*, 2013, Figure 1-5). Studies have shown that the largest proportion of suppressor alleles associated with the RNase H domain map to the  $\beta$ -hairpin region, suggesting that this region is important for the function of Prp8 in the spliceosome (Ritchie *et al.*, 2008; Yang *et al.*, 2008; Pena *et al.*, 2008).

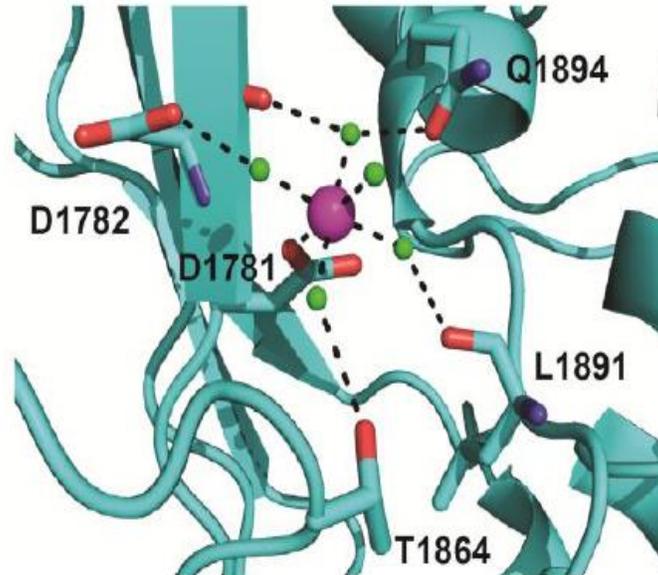
Many RNase H-like enzymes have been proposed to catalyze RNA cleavage where two divalent metal ions promote phosphodiester bond cleavage by hydrolysis. This reaction involves two metals well positioned to activate the water nucleophile and stabilize the transition state (Steitz and Steitz, 1993; Nowotny *et al.*, 2005). However, in the RNase H domain of Prp8 only one such canonical metal-binding site is spatially conserved – two aspartates D1781 and D1782 (Ritchie *et al.*, 2008). Mutation of either of these aspartate residues to alanine in yPrp8 has been shown to be lethal or severely impair growth, indicating that these metal ligands are important for splicing in yeast.

In the structure of the hPrp8 RNase H domain, a magnesium ion was observed in the canonical RNase H domain active site of the open conformation and not closed

conformation. The magnesium ion is coordinated by inner sphere contact with the side chain of Asp1781 and outer sphere coordination via five ordered water molecules to the carboxylate of D1782, backbone carbonyls of D1782 and L1891, the amide carbonyl of Q1894, and the hydroxyl of T1864 (Figure 1-6). These findings suggest that conformational change between the closed conformation and open conformation is required for metal binding to the RNase H site (Schellenberg *et al.*, 2013).



**Figure 1-5.** X-ray structure of the Prp8 RH domain. Superposition of the closed (yellow) and the open (cyan) conformations observed in the asymmetric unit (Schellenberg *et al.*, 2013).



**Figure 1-6.** Detail of coordination of Mg<sup>2+</sup> ion (purple) and its inner-sphere waters (green) in the open conformation of Prp8 RNase H domain (Schellenberg *et al.*, 2013).

It was proposed that the hydrogen bonding between the side chain of R1865 and D1782 was potentially shielding the metal binding site. The structure of the wild type and R1865A mutant hPrp8 RNase H domain were considerably similar. However, R1865A had a higher metal occupancy than the wild type structure. Also, the  $\beta$ -hairpin of open conformation, which was largely disordered in the wild-type Prp8 structure, is now visible in the R1865A structure (Schellenberg *et al.*, 2013).

Structural and functional analyses of Prp8 mutant alleles that exhibit first step allele phenotype have been shown to stabilize the closed conformation (monomer a) of Prp8 RNase H domain, implying that the closed conformation is associated with the first step of splicing. Second step alleles stabilize the metal binding, open conformation (monomer b) of the Prp8 RNase H domain, suggesting that the open conformation of the

Prp8 RNase H domain is associated with the second step of splicing (Schellenberg *et al.*, 2013).

The ~70% identity between yPrp8 and hPrp8 RNase H domains allows for the interpretation of hPrp8 structures in light of known yeast alleles.

## **1-10. Summary**

Spliceosome assembly on the pre-mRNA is a highly intricate and dynamic process. Analysis of the Prp8 RNase H domain, combined with structural and biochemical data will provide insight on the nature of the spliceosomal active site and on the conformational changes taking place during different steps along the splicing pathway. The following chapters will describe the high-resolution X-ray structures and characterization of novel mutant alleles of a potential spliceosomal active site protein, the U5 snRNP component Prp8.

## **Chapter 2**

### **Materials and Methods**

## **2-1. Materials and Methods**

### **2-1.1. Protein expression and purification**

A cDNA representing an N-terminal extension of hPrp8 1831-1990 encoding hPrp8 amino acids 1769-1990 RNaseH domain was cloned into the EcoRI and HindIII sites of pMAL-c2x plasmid (NEB) using PCR primers. TEV protease cleavage site was inserted between maltose-binding protein (MBP) and the RNaseH domain of Prp8. All Prp8 mutants were generated using the overlapping PCR method and confirmed by sequencing. The resulting MBP tagged-Prp8 protein was expressed in *E. coli* and purified by sequential amylose resin, size exclusion chromatography, TEV cleavage and anion exchange chromatography.

### **2-1.2. Crystallization**

Crystals of hPrp8 aa 1769-1990 were grown at 23°C using the hanging drop vapor diffusion technique. Protein crystals were grown by mixing 1 µl of 10 mg/ml protein solution (10 mM Tris, pH 8.0, 0.1 mM EDTA, 5 mM DTT, 0.02% NaN<sub>3</sub>) with 1 µL of precipitant (10-14% PEG 4000, 100 mM Tris pH 8.0, and 300 mM MgCl<sub>2</sub>). Crystals were cryoprotected in reservoir solution containing 15% glycerol, and frozen in liquid nitrogen for data collection (Schellenberg *et al.*, 2013).

### **2-1.3. Data collection and processing**

Prp8 mutant protein crystal data was collected at beam CMCF-1D at the Canadian Light Source, University of Saskatchewan, Saskatoon. Data were processed and scaled using the HKL package (Otwinowski and Minor, 1997).

#### **2-1.4. Model building and refinement**

The structures were solved using molecular replacement using the program Phenix MR (Adams *et al.*, 2010) with PDB id 4JK7 as a starting model. Iterative cycles of refinement in Phenix refine against the merged dataset and manual model building using COOT was used to complete and refine the models.

#### **2-1.5. Construction of RNA substrate**

The ACT1 pre-mRNA was made by *in vitro* transcription with [ $\alpha$ -<sup>32</sup>P]-ATP. The 25  $\mu$ L reaction was incubated at 37°C for 4 h, with the final concentrations of the various components as follows: 40 mM Tris pH 8.0, 6 mM MgCl<sub>2</sub>, 10 mM NaCl, 2 mM spermidine, 0.5 mM CTP, 0.5 mM GTP, 0.5 mM UTP, 60  $\mu$ M ATP, 1.32  $\mu$ M [ $\alpha$ -<sup>32</sup>P]-ATP (PerkinElmer), 5 ng DNA template, 10 mM DTT, and 1  $\mu$ L T7 RNA polymerase. The reaction was then separated on a 7%, 19:1 Acrylamide: Bis-acrylamide, 8 M Urea PAGE.

#### **2-1.6. Creation of mutant *S. cerevisiae* strains**

Mutant PRP8 genes were created by gap repair of plasmid pJU186 (Umen and Guthrie, 1995), which contains a HIS selectable marker and the PRP8 gene. Confirmed mutant plasmids were transformed into strain JDY8.06 (*ura3-52, leu2-3,-112, ade2, his3-A1, trpl-289, prp8::LEU2, pJU169 (PRP8, URA3, CEN, ARS)*), a gift from Richard Grainger and Jean Beggs, University of Edinburgh, UK, containing wild-type Prp8 on a counter-selectable URA3-marked plasmid (Brown and Beggs, 1992). After transforming the JDY8.06 cells with the pJU186 plasmid, cells were selected by growth in a medium

lacking histidine and leucine at 30°C for 16 hours. Cells lacking the URA3 plasmid were selected by streaking the transformants once on medium lacking in histidine and leucine, but containing 5-fluoroorotic acid (5-FOA) (Boeke *et al.*, 1987). Cells from a single colony that survived on 5-FOA plates were grown in media lacking histidine, and total DNA was extracted by using a DNeasy kit (Qiagen). All mutant Prp8 strains were verified by sequencing.

Copper-resistant strains were created by co-transforming the ACT1-CUP1 plasmid containing the wild-type sequence, BSC or BSG mutation at the branch site along with pJU186 plasmid containing wild-type or mutant Prp8 into yJU75 (MATa, ade2 cup1Døura3 his3 leu2 lys2 prp8DøLYS2 trp1, pJU169 (PRP8 URA3 CEN ARS) (Umen and Guthrie, 1996; Courtesy of Dr. Jonathan Staley). Cells lacking the URA3 plasmid but containing the mutant pasmid along with the ACT1-CUP1 plasmid were selected by streaking the transformants once on medium lacking in histidine and leucine, but contained 5-FOA. CUP1 plasmid has selectable marker leucine. Cells from a single colony that survived on 5-FOA plates were grown in media lacking histidine, and total DNA was extracted by using a DNeasy kit (Qiagen). All mutant Prp8 strains were verified by sequencing.

### **2-1.7. Growth Assays**

Spot test analysis was performed by inoculating the yeast strains in YPD media (1% yeast extract, 2% peptone, 2% dextrose) and growing the cells overnight at 30°C with shaking. Next day, the OD at 600nm was determined, and 10µl serial dilutions of

$1 \times 10^6$ ,  $2 \times 10^5$ ,  $4 \times 10^4$ ,  $8 \times 10^3$ ,  $1.6 \times 10^3$  cells/ml were spotted onto YPD plates. Plates were photographed after 2 days at 30°C.

For the copper growth assay, cultures containing the BSG ACT1-CUP1 and Prp8 mutant plasmids were grown overnight in media lacking leucine and histidine, diluted to  $A_{600} = 0.2$  and equal volumes were spotted onto SDC-agar plates lacking leucine and histidine, and containing 0-0.25 mM concentrations of  $\text{CuSO}_4$  (Lesser and Guthrie, 1993). Plates were photographed after 3 days at 30 °C.

### **2-1.8. Primer extension**

Cell cultures containing the ACT1-CUP1 plasmid and mutant Prp8 plasmid were grown overnight in SDC media lacking histidine and leucine, and diluted to 5 ml  $A_{600}=0.2$ . The diluted cells were grown in SDC-histidine-leucine media for an additional 6 hours to  $A_{600}=1.0$ . Yeast RNA was extracted using RNeasy mini kit (Qiagen) using manufacturer's instructions.

Primer extensions were carried out using the YAC6 primer complementary to exon 2 of CUP1 (Query and Konarska, 2004). The primer was end-labeled with [ $\gamma$ - $^{32}\text{P}$ ]-ATP (PerkinElmer). Primer extensions were performed by using the RevertAid H minus First Strand cDNA Synthesis Kit (Fermentas). A 12  $\mu\text{l}$  of the mixture (1  $\mu\text{g}$  total RNA and 2 pmol labeled primer) was heated to 70 °C for 10 minutes and slowly cooled down to 40 °C. The reaction was then chilled on ice, supplemented to 20  $\mu\text{l}$  (4  $\mu\text{l}$  reaction buffer, 1 U RNase inhibitor, 1 mM dNTP mix and 10 U reverse transcriptase), then incubated for 5 min at 37 °C and 55 min at 42 °C. The reaction was terminated, and the RNA was degraded with 0.5 M NaOH at 70 °C. Extension products were extracted with

phenol/chloroform/isoamyl alcohol and chloroform, then ethanol-precipitated. The pellet was resuspended, separated in gels containing 7% polyacrylamide and 8 M urea and visualized by autoradiography.

Quantification was performed using ImageQuant (Molecular Dynamics). The volume of each band was adjusted by subtracting the background noise. The first- and second-step efficiencies were then calculated as shown in Equation 1 and 2.

$$\text{1st step efficiency} = \frac{[\text{Lariat intermediate}] + [\text{mRNA}]}{[\text{Lariat intermediate}] + [\text{mRNA}] + [\text{pre - mRNA}]} \quad (\text{Eq.1})$$

$$\text{2nd step efficiency} = \frac{[\text{mRNA}]}{[\text{Lariat intermediate}] + [\text{mRNA}] + [\text{pre - mRNA}]} \quad (\text{Eq.2})$$

## **Chapter 3**

**Local reorganization of the Prp8 RNase H domain unmasks a conserved metal-binding site.<sup>1</sup>**

<sup>1</sup> This work was performed in collaboration with Erin Garside and Matthew Schellenberg who solved the X-ray structures.

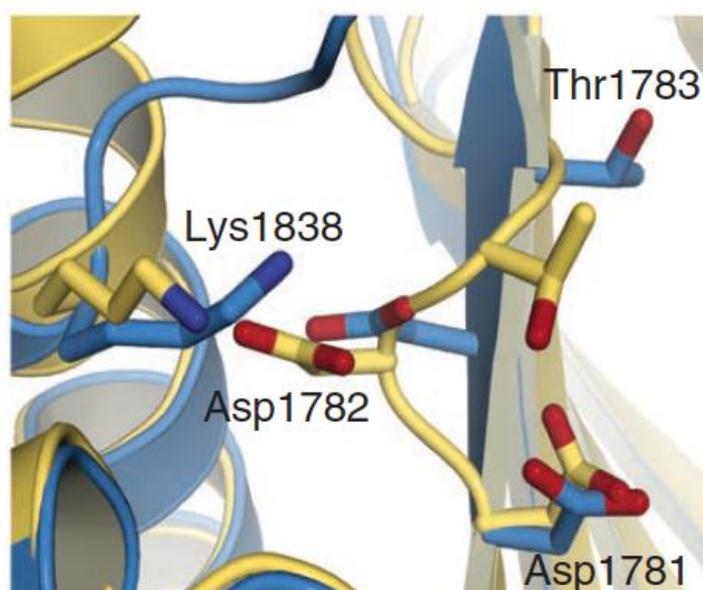
### 3-1. Introduction

The stereochemistry of splicing indicates the presence of two active sites in the spliceosome, suggesting that the spliceosome would exist in two conformations during the two transesterification steps of splicing. A previous study (Query and Konarska, 2004) has proposed a model where the spliceosome adopts two distinct conformations between the first and the second step of splicing. These two conformations are in equilibrium with each other, where two sets of Prp8 alleles, namely first step and second step alleles, favor one conformation over the other by shifting the equilibrium. A first step allele would favor the first step conformation and thus suppress a defect in the first step but enhance a defect in the second step of splicing. A second step allele would favor the second step conformation and thus suppress a defect in the second step but enhance a defect in the first step.

Two distinct conformations of the Prp8 RNase H domain have been studied crystallographically, where two monomers were observed in the asymmetric unit (Schellenberg *et al.*, 2013). Monomer a is associated with the closed conformation where no  $Mg^{2+}$  ion was observed in the putative active site. Monomer b is represented by the open conformation where  $Mg^{2+}$  coordination was observed at the conserved RNase H metal-binding site and the  $\beta$ -hairpin is disrupted to form a displaced loop.

One consequence of the structural rearrangement that takes place in the transition from the closed conformation to the open conformation is that the D1782 moves closer to D1781 and displacement of the residue T1783 upward by 4 Å (Schellenberg *et al.*, 2013, Figure 3-1). T1783 in hPrp8 is highly conserved in the RNase H domain among many organisms. In the closed conformation of the RNase H domain of hPrp8, T1783 is

shown to block the metal binding site. Displacement of T1783 in the transition of closed-to-open conformation exposes the metal-binding site in the open conformation to allow  $Mg^{2+}$  coordination. This observation led us to investigate the possible role of T1783 in the local reorganization of the Prp8 RNase H domain during the conformational change. In this chapter, I structurally and functionally characterized two new hPrp8 alleles, T1783A and T1783S (yeast T1855A and T1855S).



**Figure 3-1** Superposition of the X-ray structure of the Prp8 RNase H domain closed (yellow) and open (cyan) conformations. Displacement of T1783 upon conformational rearrangement facilitates binding of  $Mg^{2+}$  at the canonical site. The metal ion bound in the open conformation is not shown for clarity. (Schellenberg *et al.*, 2013).

## 3-2. Results

### 3-2.1. Copper growth assay

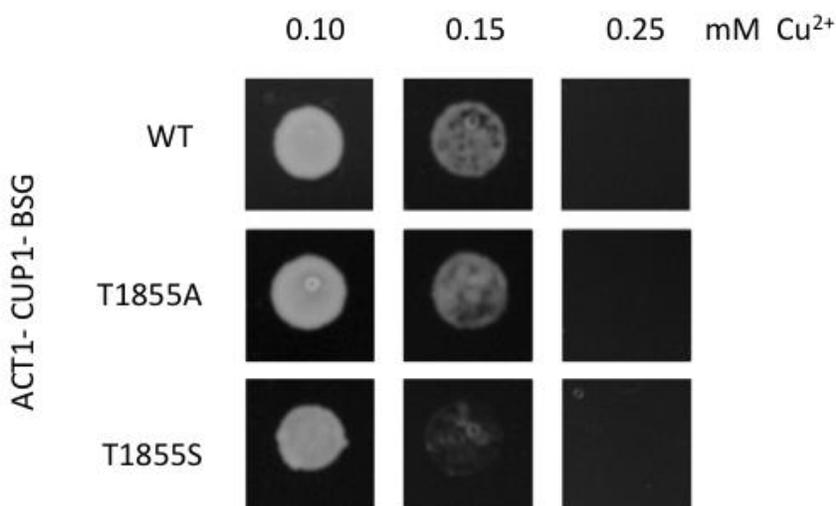
The effect of the mutations on splicing was monitored *in vivo* by using a copper-resistance reporter system called the ACT1–CUP1 reporter in which part of the ACT1 gene, including the intron, is fused in frame to the CUP1 gene. CUP1 is responsible for imparting copper resistance. The copper resistance conferred is proportional to the level of pre-mRNA splicing. Two examples of ACT1-CUP1 reporters are ACT1-CUP1-BSC and ACT1-CUP1-BSG, in which the branch-site A of the ACT1-CUP1 intron is replaced with C or G, respectively. The change in the branch site A to C causes a severe defect in both the first and the second step of splicing, whereas a change in the branch site A to G causes a defect in the second step of splicing (Burgess and Guthrie, 1993). This reporter system allows for a simple, indirect assessment of splicing activity *in vivo* (Lesser and Guthrie 1993).

Defects in ACT1-CUP1 splicing caused by the BPS mutations lead to defective CUP1 function, resulting in less growth of these strains on copper-containing plates. A second step allele would be able to suppress the defect in the second step of splicing caused by the ACT1-CUP1-BSG plasmid, allowing the cells to grow at higher concentrations of copper than the wild type or the first step allele. Thus, mutant alleles that cause less resistance to copper exhibit a first-step phenotype, whereas mutants that confer more resistance to copper exhibit a second-step phenotype.

In order to investigate the possible role of T1855 mutants in the conformational switch to allow metal binding, we generated T1855A and T1855S yPrp8 mutants within

the RNase H domain in *S. cerevisiae*, comparing the copper growth assay of the ACT1-CUP1-BSG reporter system in the above mutants to the wild-type.

For the wild type Prp8, a partial growth inhibition was observed at 0.15 mM Cu<sup>2+</sup> and a complete growth inhibition at 0.25 mM Cu<sup>2+</sup>. T1855A allele exhibited a growth defect similar to wild-type Prp8, with complete growth inhibition at 0.25 mM Cu<sup>2+</sup>. T1855S exhibited first-step allele phenotype by aggravating the growth defect of the BSG mutation, causing less resistance to copper compared the wild type. T1855S had no growth at 0.15 mM Cu<sup>2+</sup> (Figure 3-2).



**Figure 3-2.** Spot assays showing BSG ACT1-CUP1 reporter-dependent growth of yeast containing wild-type, T1855A and T1855S mutant Prp8 allele in the presence of the indicated concentration of Cu<sup>2+</sup>.

### 3-2.2. Primer extension assay

The splicing was analyzed in greater detail by a primer extension assay using ACT1-CUP1 reporters containing the wild-type intron sequence, the A-to-C mutation at

the branch site (BSC) as well as the BSG mutation (Lesser and Guthrie, 1993). Total RNA was isolated from the above strains containing the *HIS* plasmid with the desired Prp8 mutation and a co-transformed ACT1-CUP1-BSG reporter plasmid. A 5' end labeled YAC6 primer complimentary to exon 2 of ACT1-CUP1 reporter plasmid was used to carry out the primer extension. The primer anneals to the RNA and reverse transcriptase synthesizes a complementary DNA (cDNA) from the total RNA. The resulting radiolabeled cDNA products were analyzed on denaturing gel. The amount of cDNA obtained is a measure of the amount of target RNA, and the size of the cDNA reflects the distance from the primer to the 5' end or the branch of the RNA, thus enabling the analysis of levels of pre-mRNA, mRNA, and lariat intermediate.

The values obtained from the primer extension assay represent steady-state levels of RNA *in vivo*. The first step alleles which enhance the first step will have increased levels of lariat intermediate (the first step product) and reduced levels of mRNA (the second step product), while second step alleles will have decreased lariat intermediate levels and increased mRNA product because all the products of the first step are going through the first step of splicing.

All the Prp8 alleles tested completely spliced the wild-type reporter. BSC reporter is known to inhibit both steps of splicing while the BSG reporter is known to inhibit the second step of splicing. When the BSC reporter was used, the potential first-step allele T1855S showed reduced abundance of mRNA and increased abundance of lariat intermediates. The average ratio of the first step % to the second step % of T1855S-BSC allele is higher than the ratio for the wild-type, indicating that T1855S mutant is able to suppress a defect in the first step caused by the BSC reporter. As for the

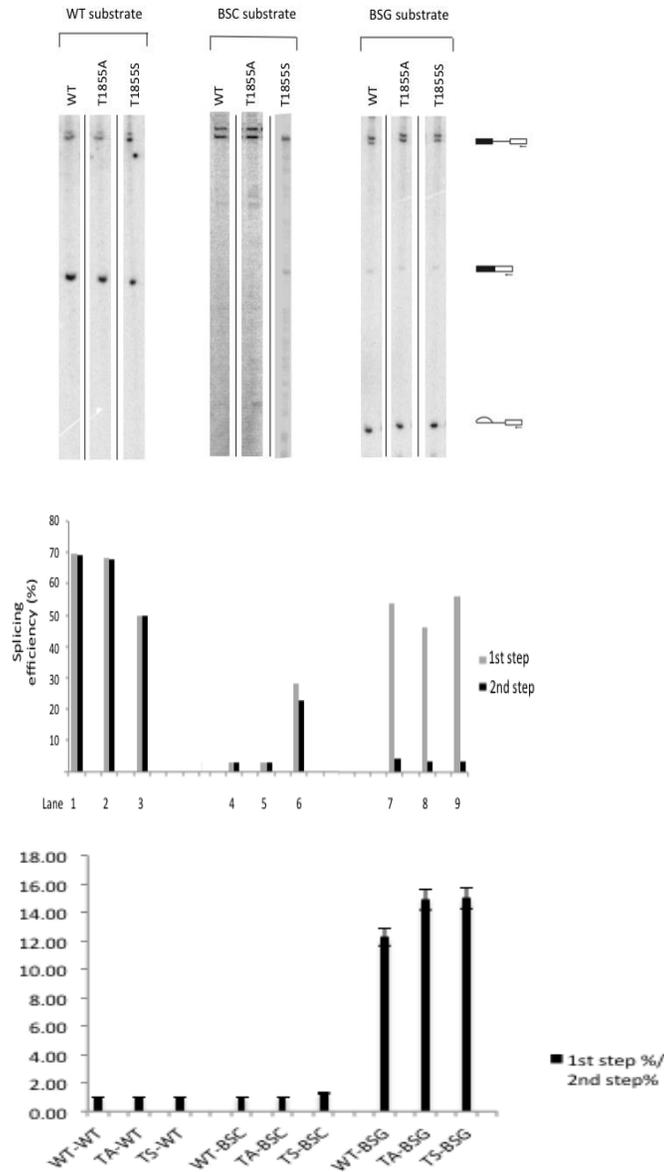
BSG reporter a decreased level of mRNA was obtained with the first step alleles. T1855A showed mRNA levels similar to the wild-type reporter (Figure 3-3). The average ratio of the first step % to the second step % of T1855S-BSG allele is higher than the ratio for the wild-type, indicating that the T1855S mutant is able to undergo the first step of splicing but is unable to suppress a defect in the second step of splicing caused by the BSG reporter. T1855A and T1855S primer extension results are consistent with the copper growth results. These results indicate that T1855S Prp8 is a first step allele while T1855A exhibits wild-type phenotype.

### 3-2.3. Structural analysis

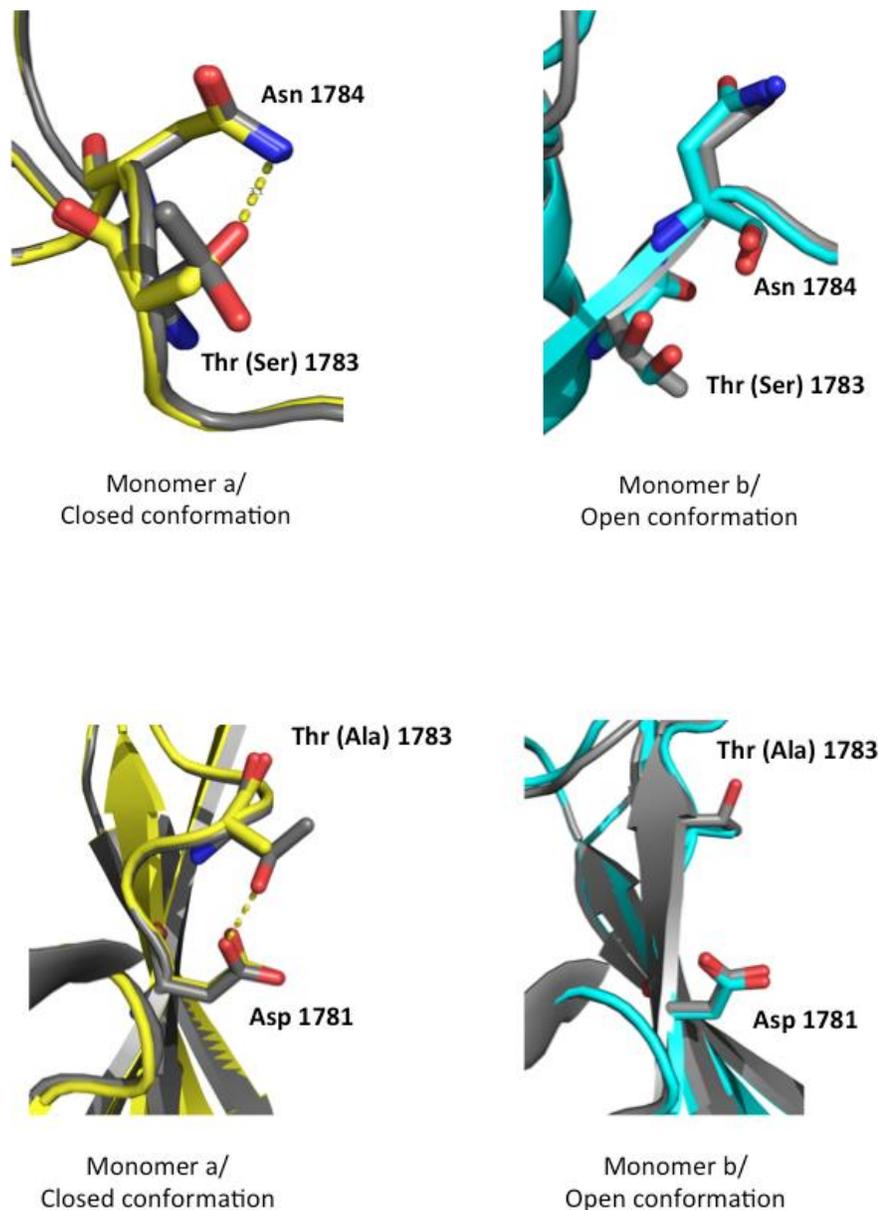
To perform structural analysis, we mutated T1783 in hPrp8 (corresponding to T1855 in yPrp8) to alanine and serine. The R1865A mutation was included because it makes the  $\beta$ -hairpin loop visible in the open conformation (Schellenberg *et al.*, 2013). We crystallized and solved the X-ray structures of T1783S/R1865A and T1783A/R1865A hPrp8 double mutants. In both T1783S/R1865A and T1783A/R1865A hPrp8 double mutants, a bound metal ion was observed at the canonical RNase H site in the open conformation and not in the closed conformation.

T1783S/R1865A hPrp8 structure revealed that the mutation results in the formation of an additional hydrogen bond to N1784 within the closed conformation. This extra hydrogen bond is effectively stabilizing the closed conformation. Open conformation accommodates the T1783S mutation with no change (Figure 3-4 A). In the T1783A/R1865A hPrp8 structure, we see that in the closed conformation, a hydrogen

bond between T1783 and D1781 is absent. Open conformation accommodates the T1783A mutation with no change (Figure 3-4 B).



**Figure 3-3.** Denaturing PAGE analysis of reverse transcriptase primer extension with  $^{32}\text{P}$ -labeled primer to examine steady-state splicing efficiencies in Prp8 mutant yeast strains. (*upper*) Primer extension analysis of RNA recovered from cells with wild-type, BSC and BSG ACT1-CUP1 plasmid and Prp8 alleles. (*middle*) Quantification of the first- and second-step efficiency. (*bottom*) Quantification of the first- and second-step efficiency performed in triplicate.



**Figure 3-4.** Characterization of Prp8 RNase H domain residue T1781 **A)** Superpositions of hPrp8 X-ray structures of wild-type (gray) with T1783S/R1865A mutant closed (yellow) or mutant open (cyan) conformers. Detail showing formation of an additional hydrogen-bonding interaction to N1784 in the closed conformation. **B)** Superpositions of hPrp8 X-ray structures of wild-type (gray) with T1783A/R1865A mutant closed (yellow) or mutant open (cyan) conformers. Detail showing loss of a hydrogen-bonding interaction to D1781 in the closed conformation.

### 3-3. Discussion

In the studies described here, we have structurally and functionally analyzed a conserved residue T1873 in the Prp8 RNase H domain. T1873 is displaced to allow metal binding by undergoing structural rearrangement in the transition of closed-to-open conformation. Functional analysis of the corresponding T1855 yeast Prp8 mutants using the copper growth assay shows that T1855S-ACT1-CUP1-BSG mutant allele caused less resistance to copper, consistent with a first-step phenotype while the T1855A-ACT1-CUP1-BSG exhibited wild-type phenotype.

We then assayed steady-state pre-mRNA splicing efficiencies of the reporter gene in these Prp8 mutants, comparing splicing of the BSC and BSG mutant ACT1-CUP1 intron using primer extension assay. BSC mutation is known to compromise both steps of splicing while the BSG mutation compromises the second step of splicing. Consistent with the results of the copper growth assay, T1855S mutant Prp8 acted a first step allele, enhancing the first step of splicing in the presence of BSC at the expense of the second step and worsening the effect of the BSG mutation. T1855A mutant Prp8 exhibited splicing efficiency similar to wild type splicing. The primer extension assay results show that T1855S is a first step allele and T1855A has wild-type phenotype.

The functional analysis results were supported by structural analysis. In the solved X-ray structures of both hPrp8 RNase H mutants, a bound metal ion was observed at the canonical RNase H site in the open conformation but was absent in the closed conformation.

In T1783S hPrp8 RNase H domain, the additional hydrogen bond between S1783 and N1784 within the closed conformation is effectively stabilizing the closed conformation relative to the open conformation. T1783 is unable to adopt this rotamer due to clashes, therefore we do not see an extra hydrogen bond in the wild-type structure. Stabilization of the closed conformation shifts the equilibrium to favor the first step of splicing. This is a characteristic of the first-step allele. Open conformation accommodates the T1783S mutation with no change in the hydrogen bonding. In T1855A Prp8 mutation, there is a loss of a hydrogen bond between T1783A and D1781 in the closed conformation.

### **3-4. Summary**

An earlier study has shown that Prp8 rearrangement unmasks a conserved metal-binding site that plays a role in the conformational change and the Mg<sup>2+</sup> coordination by Prp8 in the spliceosome. The conformational rearrangement includes displacement of T1783 in the open conformation to facilitate binding of the Mg<sup>2+</sup> at the canonical metal binding site (Schellenberg *et al.*, 2013).

In the studies described in this chapter, we have used X-ray crystallography and yeast genetic functional assays to study two new Prp8 RNase H domain mutants designed based on the conserved T1783 residue. Taken together, the studies described here provide functional and structural characterization of the two new alleles, T1783A and T1783S in the Prp8 RNase H domain.

## Chapter 4

### Local regulation of the $\beta$ -hairpin/ loop dynamics<sup>1</sup>

<sup>1</sup> This work was performed in collaboration with Erin Garside and Matthew Schellenberg who solved the X-ray structures.

#### 4-I. Introduction

The Prp8 RNase H domain has a 17 amino acid insertion within its RNase H fold. In the closed conformation these amino acids are well structured to form a two-stranded anti-parallel  $\beta$ -sheet extruding from the RNase H core, while in the open conformation this region is disordered to form a displaced loop (Ritchie *et al.*, 2008). A large number of suppressor allele of the Prp8 RNase H domain map to the  $\beta$ -hairpin region (Grainger and Beggs, 2005, Figure 4-I). In this chapter, I will discuss two separate mutant alleles that occupy the  $\beta$ -hairpin region, L1798 and N1797, within the  $\beta$ -hairpin region. We will structurally and functionally analyze how a series of mutations of the above mentioned amino acids affect the stabilization of the  $\beta$ -sheet/loop region and shift the equilibrium towards first-step or second-step allele.

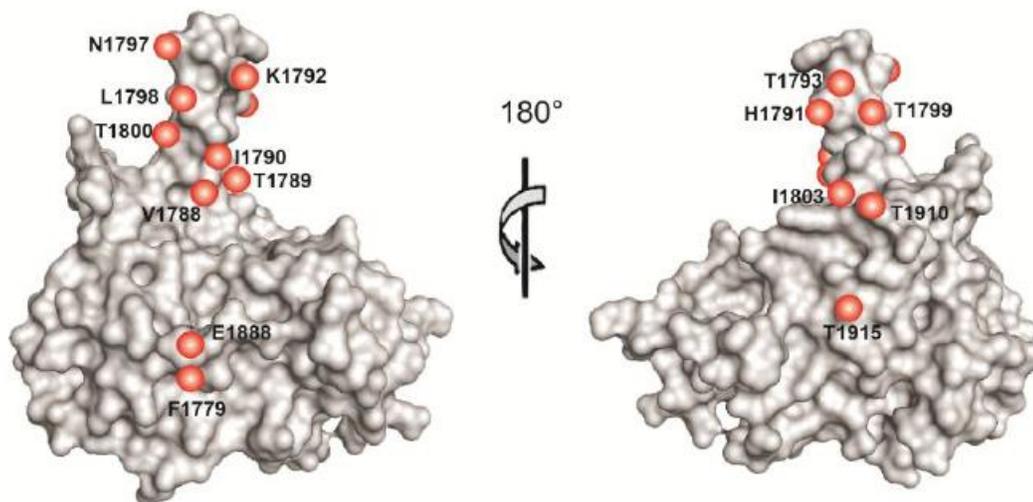


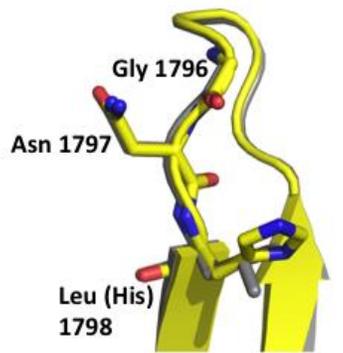
Figure 4-I. Mapping of yeast mutant alleles on the Prp8 RNase H domain. Shown are the positions of amino acids corresponding to mutant alleles (red)

The V1870N yPrp8 mutant (corresponding to the L1798N hPrp8 mutant) has been found to display a strong second-step phenotype (Query and Konarska, 2004; Liu *et al.*, 2007). To further investigate the nature of this allele, we made three other mutations in *S. cerevisiae*, namely V1870H, V1870A, and V1870D.

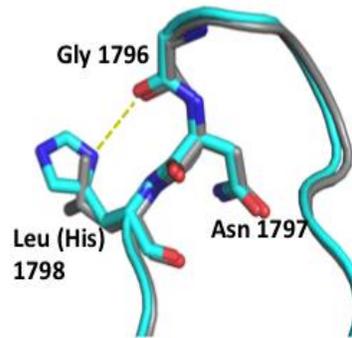
On analyzing the predicted model structure of the L1798H hPrp8, we observed that the L1798H mutant could be a potential second step allele because the imidazole ring of H1798 will be able to hydrogen bond to the carbonyl oxygen of G1796 and stabilize the open conformation (4-2 A). On the other hand, looking at the model structure of the L1798A hPrp8 mutant we predicted that L1798A would have no effect on closed or open conformations because there would be no change in hydrogen bonding relative to the wild-type structure (Figure 4-2 B). In case of L1798D hPrp8 mutant model, D1798 is not a potential hydrogen donor, so we expect no change in the hydrogen bonding in the two conformations compared to the wild-type structure (Figure 4-2 C).

On analyzing the N1797D hPrp8 RNase H domain predicted model, we can expect that the N1797D mutation could potentially stabilize the closed conformation and thus, exhibit a second-step allele phenotype.

A.

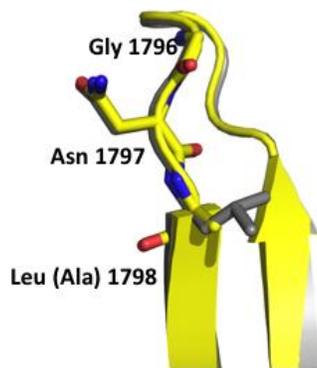


Closed conformation

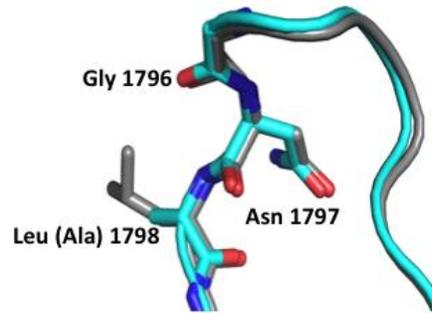


Open conformation

B.

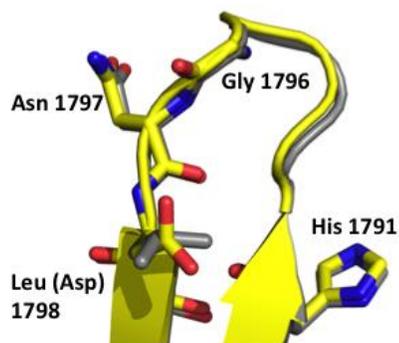


Closed conformation

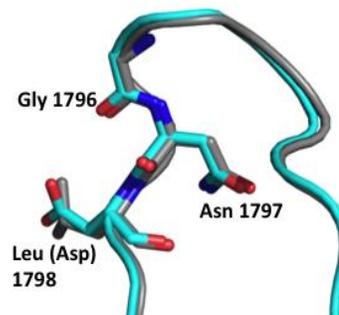


Open conformation

C.



Closed conformation



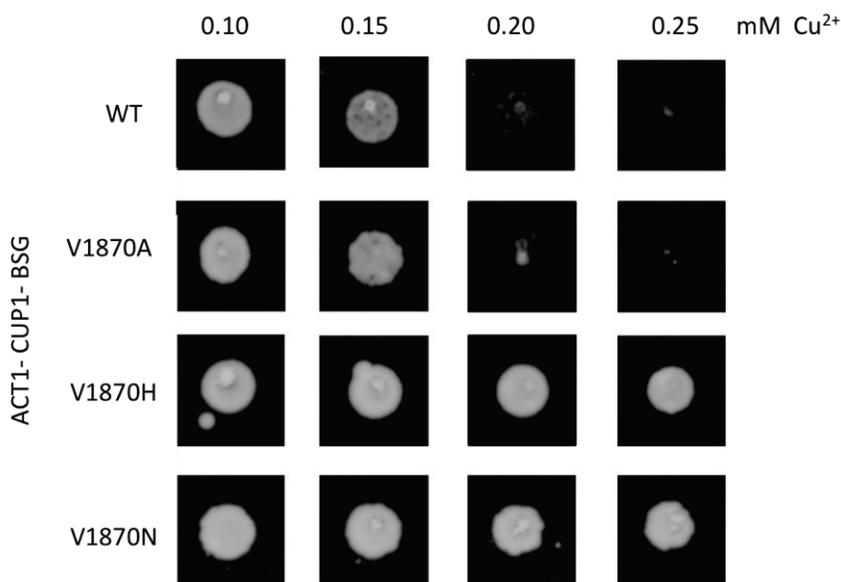
Open conformation

Figure 4-2. Models for the L1798 mutants for closed conformation (*left*) and open conformation (*right*). A) L1798H. B) L1798A. C) L1798D

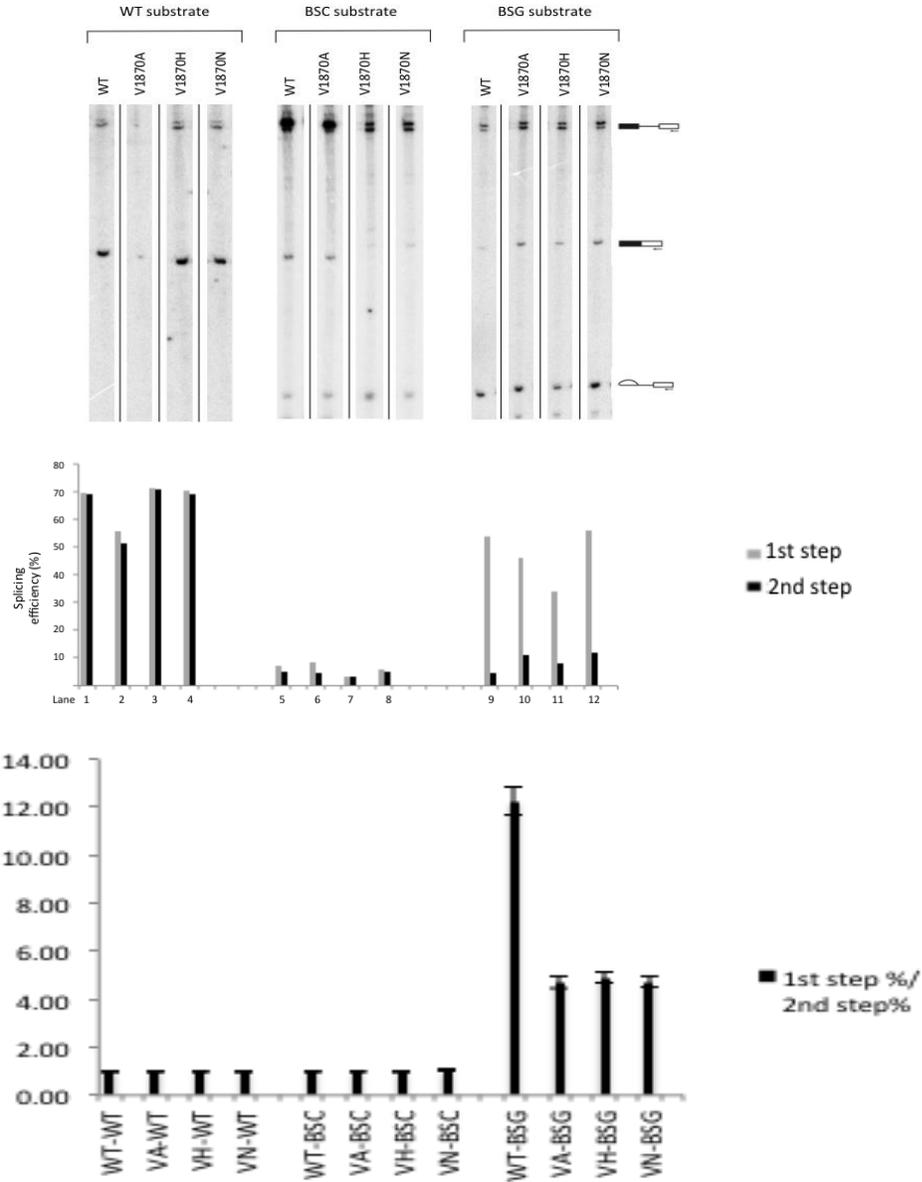
## 4-2. Results

### 4-2.1. Mutations of L1798 residue regulate the dynamics of the $\beta$ -hairpin region

If the mutation of the L1798 residue results in the rearrangement of the  $\beta$ -hairpin region, thus shifting the equilibrium towards the closed or open conformation, it would impair splicing as well as the growth of the cells. Copper growth assay in yeast shows that both V1870N and V1870H confer complete resistance to a high concentration of 0.25 mM  $\text{Cu}^{2+}$ . Copper growth results for V1870A mutation has a phenotype similar to the wild-type, with partial growth at 0.20 mM  $\text{Cu}^{2+}$  and complete growth inhibition at 0.25 mM  $\text{Cu}^{2+}$  (Figure 4-3). The V1870D mutation was lethal in yeast Prp8.



**Figure 4-3.** Spot assays showing BSG ACT1-CUP1 reporter-dependent growth of yeast containing wild-type, V1870A, V1870H and V1870N mutant Prp8 allele in the presence of the indicated concentration of  $\text{Cu}^{2+}$ .

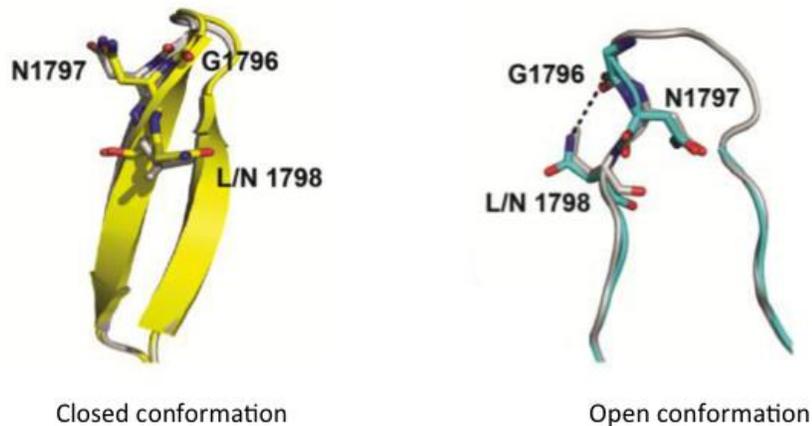


**Figure 4-4.** Denaturing PAGE analysis of reverse transcriptase primer extension with  $^{32}\text{P}$ -labeled primer to examine steady-state splicing efficiencies in Prp8 mutant yeast strains. (*upper*) Primer extension analysis of RNA recovered from cells with wild-type, BSC and BSG ACT1-CUP1 plasmid and Prp8 alleles. (*middle*) Quantification of the first- and second-step efficiency. (*bottom*) Quantification of the first- and second-step efficiency performed in triplicate.

The primer extension assay results using ACT1-CUP1-BSG reporter as a substrate for the V1870H and V1870N Prp8 mutants, shows an increased level of mRNA relative to the wild type protein. The overall splicing efficiency of the above mutants and

the wild-type Prp8 was low for the ACT1-CUP1-BSC substrate (Figure 4-5). Splicing efficiency of wild-type was impaired for both steps of splicing for the ACT1-CUP1-BSC substrate and impaired for the second step of splicing for the ACT1-CUP1-BSG substrate. The primer extension assay was done in triplicate and a graph showing error bars with a standard deviation of 1 is shown (Fig 4-4). The average ratio of the first step % to the second step % of V1870H-BSG, V1870N-BSG allele is lower than the ratio for the wild-type, indicating that V1870H and V1870N mutants are able to suppress a defect in the second step caused by the BSG reporter. The V1870A mutant, however, exhibited second step phenotype according to the primer extension assay and the average ratio of the first step % to the second step %.

In order to further investigate these potential second step alleles, we solved the X-ray structures of the corresponding mutation in the human Prp8 protein, L1798N, L1798A, and L1798H. We included the R1865A mutation in all of the above Prp8 mutant proteins because the R1865A mutation makes the loop region more ordered in the open conformation (Schellenberg *et al.*, 2013). On solving the L1798N Prp8 structure we see that the  $\beta$ -hairpin in the open conformation is disrupted and poorly ordered, making it difficult to assign the structure in this region. It is seen that the mutation L1798N forms an additional hydrogen-bond to the carbonyl oxygen of G1796 in the open conformation. The non-metal binding  $\beta$ -hairpin closed conformation accommodates the L1798N mutation with no changes in structure because the side chain of L1798 is not positioned near any hydrogen-bond donors or acceptors (Figure 4-5). This suggests that L1798N mutation stabilizes the open conformation in hPrp8 RNase H domain (Schellenberg *et al.*, 2013).



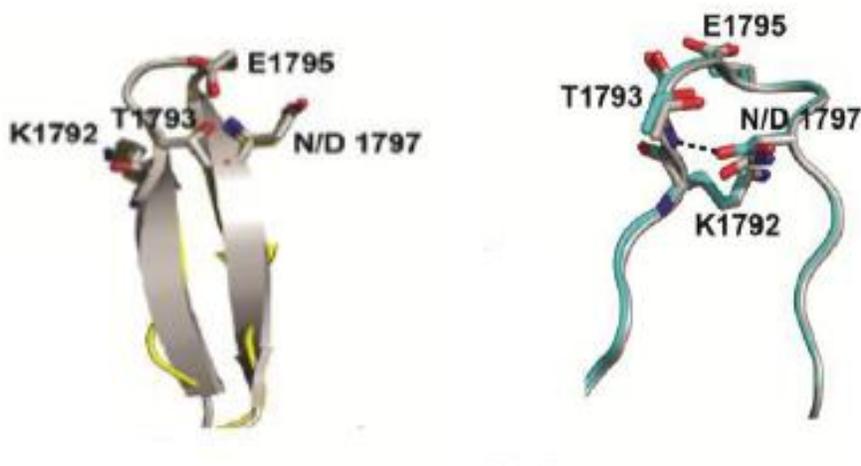
**Figure 4-5.** Details of the 2.2 Å X-ray structure of the human L1798N mutant corresponding to the yeast V1870N second-step allele. (*left*) The non-metal binding  $\beta$ -hairpin conformation of the closed conformation accommodates the mutation with no changes in structure. (*right*) The L1798N mutation forms an additional hydrogen bond to the carbonyl oxygen of G1796 in the open conformation.

Unfortunately, in the crystal structures of L1798A/R1865A and L1798H/R1865A hPrp8 RNase H mutants that we have solved, the  $\beta$ -hairpin region was largely disordered, as opposed to L1798N/R1865A where only the top portion of the  $\beta$ -hairpin in the open conformation was disordered. This makes it difficult to distinguish the closed conformation from the open conformation in the L1798A and L1798H structures.

#### 4-2.2. Mutations of N1797 residue regulates the dynamics of the $\beta$ -hairpin region

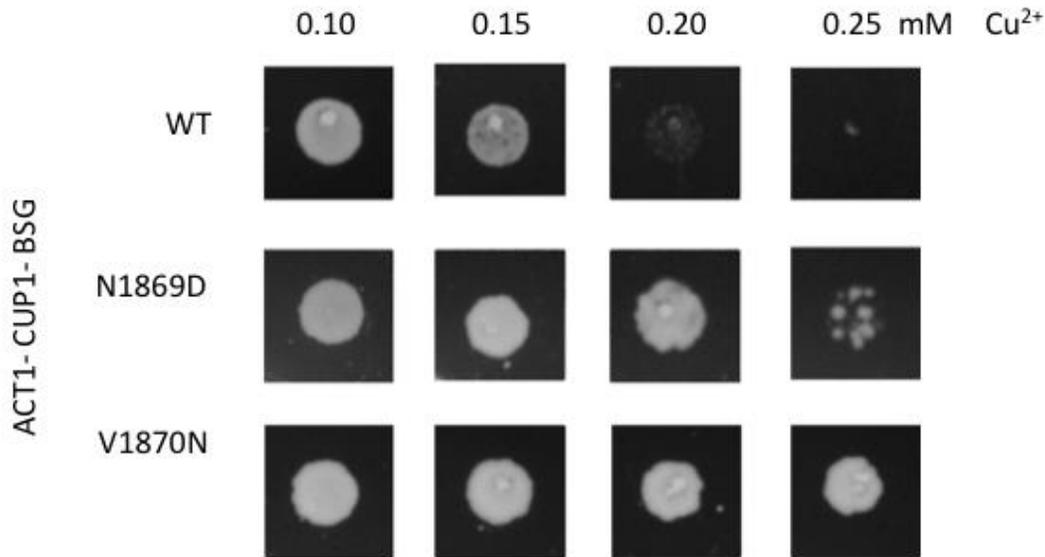
On examining the protein structure of amino acid N1797 adjacent to L1798, we see that the carbonyl of the amide group of N1797 forms a hydrogen bond with the peptide of T1793 in the open conformation. When N1797 is mutated to aspartate, the negative charge on the oxygen increases and thus the hydrogen bond is further strengthened. This results in a more ordered region in the open conformation in N1797D

relative to the wild-type protein structure. In the closed conformation, N1797 is 4.4Å away from E1795, indicating a possibility of charge-charge repulsion in this open conformation. However, in the closed conformation of N1797D structure, E1795 favors the conformation which is further away from N1797D. This suggests that a combination of the above two effects stabilizes the open conformation conformation in N1797D Prp8 mutant relative to the closed conformation conformation (Figure 4-6).



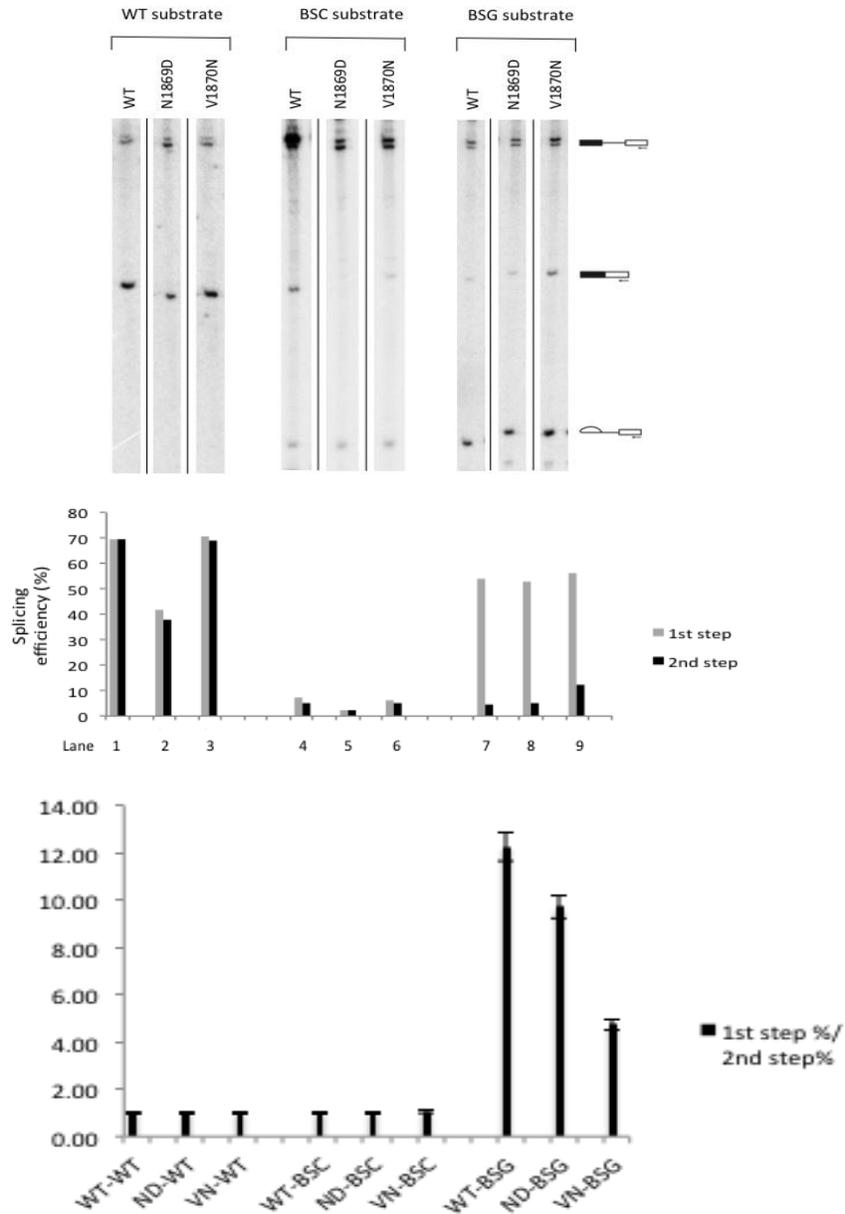
**Figure 4-6.** Details of the X-ray structure of the human N1797D mutant corresponding to the yeast N1869D second-step allele. (left) The non-metal binding  $\beta$ -hairpin conformation of closed conformation accommodates the mutation. (right) The N1797D mutation forms a hydrogen bond to the peptide nitrogen of T1793 in open conformation. This hydrogen bond is likely stronger than is observed for N1797. Consistent with this, D1797 is better positioned for this hydrogen bond than N1797.

In order to further investigate the N1797D mutation in functional studies, we made the corresponding N1869D mutation in yeast. Copper growth assay results show that N1869D confers higher resistance to copper relative to the wild-type protein but not as high as V1870N mutant which is a strong second step allele (Figure 4-7).



**Figure 4-7.** Spot assays showing BSG ACT1-CUP1 reporter–dependent growth of yeast containing wild-type, N1797D and V1870N mutant Prp8 allele in the presence of the indicated concentration of Cu<sup>2+</sup>.

Primer extension assay results support the growth assay results where we see that the growth defect of the BSG mutation is alleviated while the growth defect of the BSC mutation is aggravated. The lower average ratio of the first step % to the second step % of the N1869D allele relative to the ratio of the wild type Prp8 suggests that the N1869 mutant is able to suppress a defect in the second step of splicing caused by the BSG reporter (Figure 4-8). These results confirm that N1869D yPrp8 mutant is a mild second step allele. V1870N, which is a confirmed second step allele, was used as a positive control.



**Figure 4-8.** Denaturing PAGE analysis of reverse transcriptase primer extension with  $^{32}\text{P}$ -labeled primer to examine steady-state splicing efficiencies in Prp8 mutant yeast strains. (*upper*) Primer extension analysis of RNA recovered from cells with wild-type, BSC and BSG ACT1-CUP1 plasmid and Prp8 alleles. (*middle*) Quantification of the first- and second-step efficiency. (*bottom*) Quantification of the first- and second-step efficiency performed in triplicate.

### 4-3. Discussion

In the studies described here, using the copper growth assay and the primer extension assay, we have confirmed that V1870N yPrp8 RNase H mutant exhibits a strong second step phenotype. V1870N can grow at higher concentrations of copper because it is able to suppress the defect in the second step of splicing.

On solving the crystal structure, it was revealed that the corresponding L1798N mutation in hPrp8 introduced an additional hydrogen bond in the open conformation. L1798N forms an extra hydrogen bond to the carbonyl oxygen of G1796 and this could potentially be stabilizing the open conformation. A correlation between the sites of previously reported second-step alleles and the open conformation of the Prp8 RNase H domain indicate that the open conformation is associated during the second step of splicing in the spliceosome. In the closed conformation of L1798N, no extra hydrogen bond was observed. Thus, the results of the structural studies complement the analysis of the mutant Prp8 allele in yeast.

We also observe that V1870H is able to grow at higher concentrations of copper, which is a phenotype of a second step allele. The primer extension data analysis supports the copper growth result where the V1870H mutant is able to suppress the defect in the second step of splicing caused by the ACT1-CUP1-BSG reporter plasmid. Analysis of the solve X-ray structure of L1798H showed that only one conformation of the asymmetric unit was present where the  $\beta$ -hairpin region which distinguishes the closed conformation from the open conformation is disordered, suggesting the possibility that the structure represents a different conformation that might be an intermediate between

the closed and open conformation. On modeling the structure of the corresponding mutation in hPrp8, L1798H, we see that the imidazole ring of H1798 will be able to hydrogen bond to the carbonyl oxygen of the adjacent amino acid G1796. This additional hydrogen bond would be able to stabilize the open conformation, favoring the second step allele phenotype.

In the case of V1870A, we observe in the copper growth assay a phenotype similar to the wild-type Prp8 while the primer extension assay results show that V1870A is a second step allele. Looking at the model structure for L1798A, we can see that the alanine side chain would not be able to hydrogen bond with G1796 and thus exhibit a wild-type phenotype because of no change in hydrogen-bonding in both conformations.

The structural analysis of N1798D is consistent with the growth assay and the primer extension assay. N1798D stabilizes the open conformation by strengthening the hydrogen bond with T1793 across the loop. Functional studies support this where we can see more resistance to higher concentrations of copper as compared to the wild type. Similarly, in the primer extension assay, N1798D mutation is able to suppress a defect in the second step of splicing in the ACT1-CUP1 BSG substrate, while it enhances a defect in the splicing efficiency of the ACT1-CUP1 BSC substrate. These are characteristics of a second-step allele.

#### **4-4. Summary**

In this chapter, we have studied the local regulation of the  $\beta$ -hairpin or loop dynamics by analyzing a series of mutations of residues on the  $\beta$ -hairpin region. We observed how a mutant is able to stabilize the disordered loop, which is associated with a second step

conformation, and thus shift the equilibrium towards suppressing a defect caused in the second step of splicing.

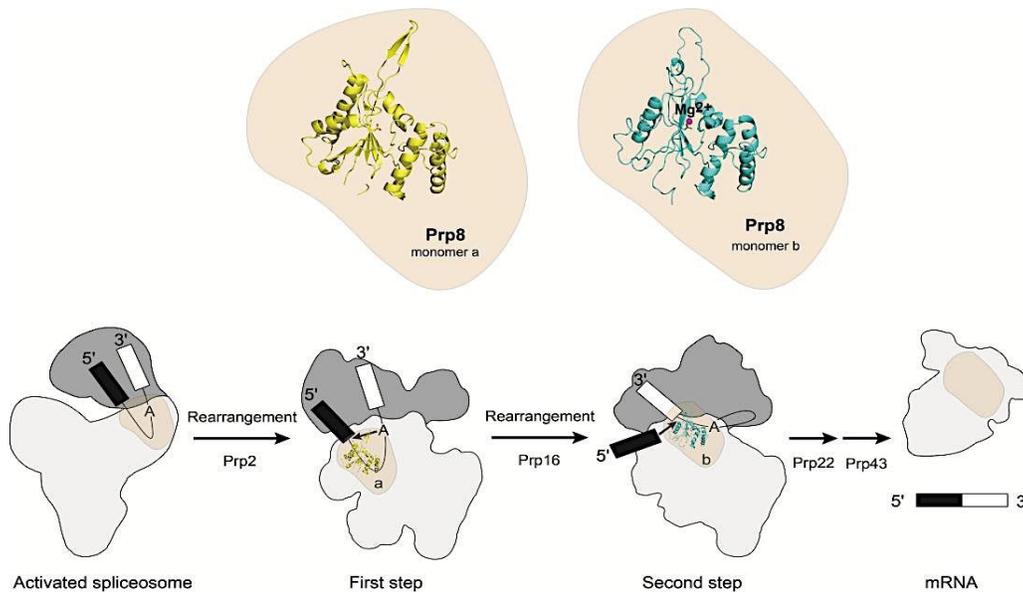
## **Chapter 5**

### **Conclusions and future directions**

## 5-1. Conclusions

The exact role of Prp8 in splicing is not fully understood. It could be stabilizing a RNA structure at the spliceosome active site, it could perform a structural role in stabilizing the second step conformation of the RNase H domain or it could have a direct role in catalysis of the second step. The unexpected discovery of an RNase H domain in Prp8 is exciting as RNase H domains are known to catalyze RNA strand cleavage via the nucleophilic attack of a water molecule on the backbone of an RNA. It also provides an insight to design additional mutants, which will not only allow further interrogation of the spliceosome, but also be useful for stalling spliceosomes at certain stages of the splicing reaction for other biochemical experiments.

In this thesis I have studied the RNase H domain of Prp8 in greater detail. We show that yeast Prp8 first- and second-step alleles favor the conformation of one of two distinct monomers observed in the crystal structure of the Prp8 RNase H domain. These data support structural and functional evidence that the RNase H domain of Prp8 undergoes a conformational switch between the first and second step of splicing. We see that magnesium ion binding is observed in the open conformation, which is associated with the second step or the exon ligation reaction of splicing (Figure 5). This supports the hypothesis that spliceosomes are ribonucleoprotein enzymes.



**Figure 5. Model representing the role of Prp8 RNase H domain in splicing.** The conformational change within the  $\beta$ -hairpin of Prp8 results in the transition of the closed conformation to the open conformation. This rearrangement results in the binding of  $Mg^{2+}$  ion (*purple*) to the active site of splicing. The shape of the complexes and the position of the Prp8 and pre-mRNA within them are all based on the EM structure studies (Boehringer *et al.*, 2004; Golas *et al.*, 2010). The U5-snRNP region is distinguished from other regions of the complex B\* and C by light grey. (Schellenberg *et al.*, 2013)

The conformational change unmasks a metal-binding site by the displacement of T1783 that is important for the catalysis of the second step of splicing. Mutation of this conserved residue T1783S resulted in a change in the hydrogen bonding, leading to the impairment in the second-step of splicing. Thus, we can see that displacement of T1783 is functionally important for the metal-binding and the transition of the closed conformation to the open conformation.

The conformational switch coupled to the binding of a functionally important metal ion implicates it as a key regulatory mechanism in promotion of the second chemical step of splicing. The precise role of the bound metal ion within the open conformation of the Prp8 RNase H domain remains to be determined and will require further structural and functional analysis.

We also looked into the importance of Prp8 RNase H mutants stabilizing the  $\beta$ -hairpin, which is a characteristic of a first step allele versus stabilizing the displaced loop, which is a characteristic of a second step allele.

The aim of the research described here is to further understand the role of Prp8 in splicing. This work will provide key structural and functional information with regard to a biochemical machine that regulates the expression of almost all protein-coding genes. Because RNA modifying reactions and protein-RNA complexes are involved in diverse processes in the cell, this work will provide additional insights to a wide spectrum of human biology.

## **5-2. Future Directions**

### **5-2.1. Designing new mutant alleles for the RNase H domain**

The conformational change of Prp8 led to the assessment of a number of mutants, which can be classified as first- or second-step alleles. As seen in chapter 3, V1870N is a strong second step allele. It also provides a starting point for the design of additional mutants, which will not only allow further interrogation of the spliceosome, but also be useful for stalling spliceosomes at certain stages of the splicing reaction for other

biochemical experiments. As an initial trial, we designed three mutants corresponding to V1870N, namely V1870A, V1870H and V1870D. We have discussed mutants V1870A and V1870H in chapter 3.

It was seen that V1870D is a lethal mutation. This was observed when the transformation of V1870D Prp8 mutant plasmid into yeast cells and growing them on 5-FOA plates killed the strain, indicating that this mutation is lethal in yeast. Looking at the predicted L1798D structure in hPrp8, we observe that the mutation results in the removal of a hydrogen donor that is proposed in L1798N mutation, indicating that there should not be any change in the hydrogen bonding between L1798D and G1796 relative to the wild-type structure. It is unclear how this change in hydrogen bonding is proving to be lethal.

To study this lethal mutation in detail in yeast, the V1870D Prp8 mutant will be expressed in the pseudodiploid strain containing a wild type, untagged copy of Prp8 on a *URA3* plasmid and a copy of a Protein A-tagged Prp8 on a *HIS* plasmid harboring the desired Prp8 mutation. The viability of the pseudodiploid yeast strains will be maintained by the wild type untagged Prp8 copy, allowing us to assess splicing efficiency of the lethal mutants. The splicing extract prepared from the pseudodiploid strain will be used to perform *in vitro* splicing followed by a pull down assay with IgG beads that have affinity for Protein A tagged Prp8 and the spliceosome associated with it.

### **5-2.2. Observe metal binding in Prp8 RNase H domain of *S. cerevisiae***

As mentioned earlier, Q1894 in hPrp8 coordinates  $Mg^{2+}$  ion via water. However, in *S. cerevisiae* serine is present at this position. This raises the question about how yPrp8 coordinates the metal ion. hPrp8 and yPrp8 share 61% sequence identity while their RNase H domains share ~70% sequence identity. Two independent groups have reported the crystal structure of yPrp8 RNase H domain, where they see only monomer a in the asymmetric unit. None of the yPrp8 RNase H domain crystal structures reported have observed metal binding (Yang *et al*, 2008; Pena *et al*, 2008). Thus, future work should focus on expressing the yPrp8 RNase H domain construct that corresponds to our hPrp8 RNase domain and try to optimize the crystallization conditions to observe both monomers and the metal binding in the canonical metal binding site of RNase H domain.

### **5-3. Summary**

Gene expression is a complex network involving multiple interactions between pre-mRNA, snRNPs and splicing factor proteins. Mutations or any disruption in the regulation of splicing can be direct causative agents of a disease or can contribute to the determinants of disease susceptibility. A large number of splicing-related diseases have been reported. A deeper understanding of the splicing pathway and the factors involved in the splicing pathway will allow accurate predictions of the changes affecting splicing and the identification of disease where disrupted splicing is the primary defect.

## References

- Adams, P.D.; Afonine, P.V.; Bunkoczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; McCoy, A.J.; Moriarty, N.W.; Oeffner, R.; Read, R.J.; Richardson, D.C.; Richardson, J.S.; Terwilliger, T.C.; Zwart, P.H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica, Section D: Biological Crystallography*, *Acta Cryst. D* **66**, 213–221
- Barash, Y., Calarco, J.A., Gao, W., Pan, Q., Wang, X., Shai, O., Blencowe, B.J., and Frey, B.J. (2010). Deciphering the splicing code. *Nature* **465**, 53-59.
- Berget, S. M., Moore, C., & Sharp, P. A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Reviews in medical virology*, *10*(6), 356–62.
- Boeke, J.D., Trueheart, J., Natsoulis, G., and Fink, G.R. (1987). 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol.* **154**, 164-175.
- Boon, K.-L., Grainger, R.J., Ehsani, P., Barrass, J.D., Auchynnikava, T., Inglehearn, C.F., and Beggs, J.D. (2007). prp8 mutations that cause human retinitis pigmentosa lead to a U5 snRNP maturation defect in yeast. *Nat. Struct. Mol. Biol.* **14**, 1077-1083.
- Brow, D.A. (2002). Allosteric cascade of spliceosome activation. *Annu. Rev. Genet.* **36**, 333-360.
- Brown, J.D., and Beggs, J.D. (1992). Roles of PRP8 protein in the assembly of splicing complexes. *EMBO J.* **11**, 3721-3729.

Calvin, K., and Li, H. (2008). RNA-splicing endonuclease structure and function. *Cell. Mol. Life Sci.* **65**, 1176-1185.

Cech, T.R. (1986). The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* **44**, 207-210.

Chan, S.P., Kao, D.I., Tsai, W.Y., and Cheng, S.C. (2003). The Prp19p associated complex in spliceosome activation. *Science* **302**, 279-282.

Chow, L. T., Gelinas, R. E., & Roberts, R. J. (1977). An Amazing Sequence Arrangement at the 5' Ends of Adenovirus 2 Messenger RNA. *Cell* **12**, 1-8.

Grainger, R.J., and Beggs, J.D. (2005). Prp8 protein: at the heart of the spliceosome. *RNA* **11**, 533-557.

Kattah, N.H., Kattah, M.G., and Utz, P.J. (2010). The U1-snRNP complex: structural properties relating to autoimmune pathogenesis in rheumatic diseases. *Immunol. Rev.* **233**, 126-145.

Koodathingal, P., Novak, T., Piccirilli, J.A., and Staley, J.P. (2010). The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. *Mol. Cell* **39**, 385-395.

Lardelli, R.M., Thompson, J.X., Yates, J.R., and Stevens, S.W. (2010). Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. *RNA* **16**, 516-528.

Lesser, C.F., and Guthrie, C. (1993). Mutational analysis of pre-mRNA splicing in *Saccharomyces cerevisiae* using a sensitive new reporter gene, CUP1. *Genetics* **133**, 851-863.

MacMillan, A.M., Query, C.C., Allerson, C.R., Chen, S., Verdine, G.L., and Sharp, P.A. (1994). Dynamic association of proteins with the pre-mRNA branch region. *Genes Dev.* **8**, 3008-3020.

Madhani, H.D., and Guthrie, C. (1994). Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.* **28**, 1-26.

Merendino, L., Guth, S., Bilbao, D., Martinez, C., and Valcarcel, J. (1999). Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature* **402**, 838-841.

Moore, M., Query, C., Sharp, P. (1993) Splicing of Precursors to mRNA by the Spliceosome. *In The RNA World*. Cold Spring Harbor Laboratory Press, pp. 303-357.

Nielsen, H., and Johansen, S.D. (2009). Group I introns: Moving in new directions. *RNA Biol.* **6**, 375-383.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.

Pandit, S., Lynn, B., and Rymond, B.C. (2006). Inhibition of a spliceosome turnover pathway suppresses splicing defects. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 13700-13705.

Pena, V., Rozov, A., Fabrizio, P., Luhrmann, R. & Wahl, M.C. (2008) Structure and function of an RNase H domain at the heart of the spliceosome. *EMBO J.* **27**, 2929–2940.

Query, C. C., Moore, M. J., & Sharp, P. A. (1994). Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes & development*, **8**(5), 587–97.

Query C.C., Strobel S.A., Sharp P.A. (1996) Three recognition events at the branch-site adenine. *EMBO J.* **15**, 1392-402.

Query, C.C., and Konarska, M.M. (2004). Suppression of multiple substrate mutations by spliceosomal prp8 alleles suggests functional correlations with ribosomal ambiguity mutants. *Mol. Cell* **14**, 343-354.

Rio, D.C. (1993). Splicing of pre-mRNA: mechanism, regulation and role in development. *Curr. Opin. Genet. Dev.* **3**, 574-584.

Ritchie, D.B., Schellenberg, M.J., Gesner, E.M., Raithatha, S.A., Stuart, D.T., and MacMillan, A.M. (2008). Structural elucidation of a PRP8 core domain from the heart of the spliceosome. *Nat. Struct. Mol. Biol.* **15**, 1199-1205.

Schellenberg, M.J., Wi, T., Ritchie, D.B., Fica, S., Staley, J.P., Atta, K.A., LaPointe, P. and MacMillan, A.M. (2013). A conformational switch in PRP8 mediates metal ion coordination that promotes pre-mRNA exon ligation. *Nat. Struct. Mol. Biol.* **20** 728-734.

Schwer, B. (2008). A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol. Cell* **30**, 743-754.

Small E.C., Leggett S.R., Winans A.A., Staley J.P. The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase. *Mol. Cell.* 2006;23:389–399.

Staley, J. P., & Guthrie, C. (1999). An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p.

Umen, J.G., and Guthrie, C. (1995). A novel role for a U5 snRNP protein in 3' splice site selection. *Genes Dev.* **9**, 855-868.

Umen, J.G., and Guthrie, C. (1996). Mutagenesis of the yeast gene PRP8 reveals domains governing the specificity and fidelity of 3' splice site selection. *Genetics* **143**, 723-739.

Valadkhan, S., Mohammadi, A., Wachtel, C., and Manley, J.L. (2007). Protein-free spliceosomal snRNAs catalyze a reaction that resembles the first step of splicing. *RNA* **13**, 2300-2311.

Valadkhan, S., Mohammadi, A., Jaladat, Y., and Geisler, S. (2009). Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11901-11906.

Valadkhan, S., and Jaladat, Y. (2010). The spliceosomal proteome: At the heart of the largest cellular ribonucleoprotein machine. *Proteomics* **10**, 4128-4141.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* **136**, 701-718.

Will, C. L., & Lührmann, R. (2001a). Spliceosomal UsnRNP biogenesis, structure and function. *Current opinion in cell biology*, 13(3), 290–301.

Wu, S., Romfo, C.M., Nilsen, T.W., and Green, M.R. (1999). Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* **402**, 832-835.

Yang, K., Zhang, L., Xu, T., Heroux, A. & Zhao, R. (2008) Crystal structure of the  $\beta$ -finger domain of Prp8 reveals analogy to ribosomal proteins. *Proc. Natl. Acad. Sci. USA* **105**, 13817–13822

Zhang, D., and Rosbash, M. (1999). Identification of eight proteins that crosslink to pre-mRNA in the yeast commitment complex. *Genes Dev.* **13**, 581–592.

Zorio, D.A., and Blumenthal, T. (1999). Both subunits of U2AF recognize the 3' splice site in *Caenorhabditis elegans*. *Nature* **402**, 835-838.